

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 07:00:00 ; Search time 3215 Seconds
(Without alignments)
8560.021 Million cell updates/sec

Title: US-10-776-213-2

Perfect score: 723

Sequence: 1 ctctcgatcagcagcacac.....aaaagacataataacat 723

Scoring table: IDENTITY_NUC

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_esc4:*
5: gb_esc5:*
6: gb_esc6:*
7: gb_esc7:*
8: gb_esc8:*
9: gb_esc9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	249	34.4	434	8	A0874584 V11B3 mt
C 2	140.8	19.5	927	9	CNS06GVN
C 3	45	6.2	821	9	CNS06GVN
C 4	42.8	5.9	1055	8	CC252581
C 5	42	5.8	1101	9	CNS017KX
C 6	41.6	5.8	1101	9	CNS012J2
C 7	41.4	5.7	436	4	BM274028
C 8	41.4	5.7	1101	9	CNS0100X
C 9	40.8	5.6	791	9	AG532718
C 10	40.4	5.6	698	8	BH527702
C 11	40.4	5.6	791	4	BI255759
C 12	40.2	5.5	1000	9	CNS02UHV
C 13	40	5.5	366	4	BG628729
C 14	40	5.5	518	1	AU077905
C 15	40	5.5	845	8	AZ541311
C 16	40	5.5	871	8	AZ674746
C 17	39.6	5.5	179	2	AM477009
C 18	39.6	5.5	665	6	CA126956
C 19	39.6	5.5	668	9	CE377922
C 20	39.6	5.5	677	9	CL814372
C 21	39.6	5.5	716	9	CL814356
C 22	39.6	5.5	1332	3	CR675673
C 23	39.4	5.4	641	9	CE491438
C 24	39.2	5.4	397	5	BQ399446

C 25	39	5.4	801	7	CK461301
C 26	39	5.4	1233	8	CC192318
C 27	38.8	5.4	A0752307	HS 5565	B
C 28	38.6	5.3	553	7	CN385766
C 29	38.6	5.3	630	8	BI9521
C 30	38.6	5.3	677	9	CE498419
C 31	38.6	5.3	1081	9	CL510366
C 32	38.6	5.3	1101	9	CNS00370
C 33	38.4	5.3	243	2	AM633159
C 34	38.4	5.3	594	8	A0383493
C 35	38.4	5.3	739	9	CE806390
C 36	38.4	5.3	1316	9	AG435289
C 37	38.2	5.3	398	2	AM133335
C 38	38.2	5.3	674	9	AG157962
C 39	38	5.3	600	8	B2312600
C 40	38	5.3	600	8	B2327427
C 41	38	5.3	691	9	CL951462
C 42	38	5.3	802	9	AG468867
C 43	38	5.3	914	8	AZ547087
C 44	38	5.3	931	8	AZ679612
C 45	38	5.3	2639	3	AF289590

ALIGNMENTS

RESULT 1
A0874584/c 434 bp DNA linear GSS 08-NOV-1999
LOCUS V11B3 mtN-3xHA/JacZ Insertion library, strain Y2278 Saccharomyces cerevisiae genomic 5', genomic survey sequence.

ACCESSION A0874584
VERSION A0874584.1 GI:6286828
KEYWORDS GSS.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE

1 (bases 1 to 434)
Rose-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., deBergues, S.A., Cheung, K.-H., Sheehan, A., Symonatsis, D., Jansen, R., Umansky, L., Heideman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G.S., and Snyder, M.
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption
Unpublished (1999)

JOURNAL

COMMENT
Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumare@yale.edu
te of mtN-3xHA/JacZ insertion.
Seq primer: GGCCTCTTCTTGAGACTAC
Class: transposon-tagged.

FEATURES

source
1..434
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="Y2278 - S288C background, clr(0) rho(0)"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mtN-3xHA/JacZ Insertion library, strain Y2278"
/note="Vector: pHS56-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS56-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mtN-3xHA/JacZ minitransposon containing lacZ, URA3, and tetr resistance."

ORIGIN

Query Match 34.4%; Score 249; DB 8; Length 434;
 Best Local Similarity 100.0%; Pred. No. 6.2e-60;
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTGATTTAGACCGACACATCATGATGCGTCATAAATAATACATACGGAATA 60
 DB 311 CTTTGATTTAGACCGACACATCATGATGCGTCATAAATAATACATACGGAATA 252

QY 61 ACCATTAAGAGCAAGACGATACCTACTTGTGAAGAAAGAGACCGCTTGTAAAGGGGAT 120
 DB 251 ACCATTAAGAGCAAGACGATACCTACTTGTGAAGAAAGAGACCGCTTGTAAAGGGGAT 192

QY 121 GGGGGCTAAGAGTCACTTCACTTTCTTTCCTCGCGGTCGGAGACCGGACCCCTCT 180
 DB 191 GGGGGCTAAGAGTCACTTCACTTTCTTTCCTCGCGGTCGGAGACCCCTCTCT 132

QY 181 CTCCTCCGACGATTTCTTCTTTCATATGCTTCTTATTCATCCGCTGTGAAGCAAC 240
 DB 131 CTCCTCCGACGATTTCTTCTTTCATATGCTTCTTATTCATCCGCTGTGAAGCAAC 72

QY 241 GCACATATGA 249
 DB 71 GCACATATGA 63

RESULT 2
 CDS06GVN 927 bp DNA linear GSS 30-NOV-2001
 LOCUS T3 end of clone AS0A00706 of library AS0A from strain CLIB 533
 DEFINITION of Saccharomyces bayanus, genomic survey sequence.
 ACCESSION AL398217 GI:12151528
 VERSION AL398217.1 GI:12151528
 KEYWORDS GSS.
 SOURCE Saccharomyces bayanus
 ORGANISM Saccharomyces bayanus
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 REFERENCE 1 (bases 1 to 927)
 AUTHORS Soulier,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
 Bolland,Fukuhara,M., Bon,E., Brotier,P., Casaregola,S.,
 de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
 Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
 Sautin,M., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
 Wincker,P. and Weissenbach,J.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 Yeast species for molecular evolution studies
 JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
 MEDLINE 20584711
 PUBMED 11152876
 JOURNAL 2 (bases 1 to 927)
 MEDLINE Bon,E., Neuvéglise,C., Casaregola,S., Artiguenave,F., Wincker,P.,
 Aigle,M. and Durrens,P.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 5.
 Saccharomyces bayanus var. uvarum
 JOURNAL FEBS Lett. 487 (1), 37-41 (2000)
 MEDLINE 20584715
 PUBMED 11152880
 JOURNAL 3 (bases 1 to 927)
 MEDLINE Direct Submission
 TITLE Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,
 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
 segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT This GSS is part of a random genomic sequencing program of thirteen
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
 exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
 lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbicola,
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
 5 kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of
 the other extremity of this insert.

FEATURES
 source
 Location/Qualifiers
 1..927
 /organism="Saccharomyces bayanus"
 /mol_type="genomic DNA"
 /strain="CLIB 533"
 /variety="uvarum"
 /db_xref="taxon:4931"
 /clone="AS0A00706"
 /clone_1ib="AS0AA"
 /note="end = T3"
 <3..>107
 /note="similar to Saccharomyces cerevisiae ORF YMR251w [
 strong similarity to YK076w and YG154c]"
 /evidence=not_experimental

misc_feature
 ORIGIN

Query Match 19.5%; Score 140.8; DB 9; Length 927;
 Best Local Similarity 69.2%; Pred. No. 6.7e-29;
 Matches 321; Conservative 0; Mismatches 107; Indels 36; Gaps 8;

QY 114 GGGGGATGAGGGCTAAGAGTCACTTCTTTCCTTCGCGGTCGGACCGGAG 173
 DB 240 GCGAGAGAGGGGGTGAAGTCACTCAC-TCTCTTCCCTTTATATGTCGCGAAG 298

QY 174 CCTCTCTCCCGCAGATTTCTTCTTCAATCTTCTTTATTCCTTAATCCGTTGA 233
 DB 299 CCCCCCTTCCCGCAGCGTTTCTGTTCT-----TTCTTTTCCCTTGG 345

QY 224 AGCAACGCACTATGATCAATATGTCGACATCTCCATGAGCTGTGATGTGAT 293
 DB 346 GACAAACGCAATGAGCTTAAGAGACCTGCTTCTTATGCTGAT-----TGTGCCA 399

QY 294 TCACAGTGTAAACGACCGTGGCTCGGAAACGTTCCCTGCTGACATTTAGAAACAG 353
 DB 400 AAACGCAATAGCGACACATGCGGCCGGAAC--GTTCTCTAATGTTTCTAAGAACG 457

QY 354 GCGTCACTTCGATAT-AGAATATATAGCGCATTTTGTAGCGCCGCGCGC----- 406
 DB 458 GCGTATGCGCCCTGCAATGATATATATGCGCATTTTGTAGCGCCGCGTGTGCTG 517

QY 407 -----GGCGCGGTTCCCAATAGGAGGCGAGTTTATCGCGAGCTCTACTTCTCT 461
 DB 518 CGCAGCGCCCTTCTCTTAATATGAGGCGCATTTATCGACATGCTTAC-TCTTCT 576

QY 462 ATTGGGTAAACCCCTTCTGTTTTCGCGCAGTGTGCTGACAGCTGCGCGAGAAC 521
 DB 577 ATTGGGAAACCCCTTCTGTTTTCGCGCAGAGGCTGACAGCTGCGCGAGATA 636

QY 522 TAGTGATTAAGAGATTAAC-TTTCGATGAGAAATTAGCAAGCG 564
 DB 637 GAGCGATTAAGAGATGTGACTTTTCATGAGAGACTTACAGAGAG 680

RESULT 3
 CDS0090X/c 821 bp DNA linear GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence T873 end of BAC #
 DEFINITION BACR19021 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL052985 GI:4934433
 VERSION AL052985.1
 KEYWORDS Drosophila melanogaster (fruit fly)
 SOURCE Drosophila melanogaster
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 821)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

SOURCE	Brassica oleracea
ORGANISM	Brassica oleracea
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euteroids II; Brassicales; Brassicaceae; Brassica.
AUTHORS	1 (bases 1 to 698)
TITLE	Tom,C.D., Van Aken,S., Uteirack,T., Koo,H. and Fraser,C.M.
JOURNAL	Whole genome shotgun sequencing of Brassica oleracea unpublished (2001)
COMMENT	Other_GSSs: B0GTD44TR Contact: Chris Town TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TP Class: sheared ends.
FEATURES	location/Qualifiers
SOURCE	1..698 /organism="Brassica oleracea" /mol_type="genomic DNA" /strain="T01000DH3" /db_xref="taxon:3712" /clone="B0GTD44" /clone_11b="B0GT" /note="Vector: pHOS1, Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"
ORIGIN	
Query Match	5.6%; Score 40.4; DB 8; Length 698;
Best Local Similarity	50.0%; Pred. No. 3.5;
Matches 101; Conservative	0; Mismatches 101; Indels 0; Gaps 0;
QY	518 AACATGATGAAGGATGATGTAACCTTCATGAGAGAAATAGCAACGAGAAAAAATGAG 577
Db	485 AAAAAATTGTTAATCTAATCGTTTGTGAGAGATTAGTACATGATAGTAAATTTGTTA 544
QY	578 GCTAGCTGGAGATTGTTTTCATCATATMAAAGGAGAAATGTTGCTCACTATGTGAC 637
Db	545 ACTAAGTGAATATTATTTGGACATATATTGATTTTGTAGTTTGGGTTTAAATATTA 604
QY	638 AGTTCTGGAGCGTCTTACTTTTATTCGACAGACATCAATATATACAGATATTTGCA 697
Db	605 AATTAGTTGTATGATTTATCCAAATTTGATTTGATTTAAAAAACAATTTTAA 664
QY	698 AAAAAAAGACTAATATTA 719
Db	665 AAAAAATACATCTAATATTA 686
RESULT 11	
BI255759/c	
LOCUS	BI255759 791 bp mRNA linear EST 17-JUL-2001
DEFINITION	602977180F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5122410 5',
ACCESSION	BI255759
VERSION	BI255759.1
KEYWORDS	GI:14809497
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: csabos-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Incyte Genomics, Inc.

FEATURES	SOURCE
<p> Cloning sequencing by: Inceye Genomics, Inc. DNA distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11298 Row: 0 Column: 19 High quality sequence start: 2 High quality sequence stop: 711. Location/Qualifiers 1. 791 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5122410" /tissue_type="cervical carcinoma cell line" /lab_host="DH10B" /clone_1b="NIH_MGC_12" /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 kb. Library prepared by Life Technologies." </p>	<p> Query Match Best Local Similarity 58.2%; Pred. No. 3.6; Matches 71; Conservative 0; Mismatches 51; Indels 0; Gaps 0; </p>
<p> 588 AGTTGTTTTCATCATATATAAAGGAGAAATTTGCTCAGCATATGACAGATTCTGGG 647 505 ACTTCTCTCACACTCATTAAGTCAGTGCAGATTGTCAGAGTAATTCACAAAGTCAGGG 446 648 ACGCTTAACTTTTATTTGTCAGAGAGCATGCAATTCATACAGATATGTCAGAAAAA 707 Db 445 ATTCTTAATTTTGTGTTGTCAGATTATCTGAGATTAAGTTTAAAGAAAAA 386 QY 708 AG 709 Db 385 TG 384 </p>	<p> RESULT 12 CDS020UH LOCUS DEFINITION 1000 bp DNA linear GSS 01-SEP-2000 Tetradon nigroviridis genome survey sequence T7 end of clone 167C22 of library G from Tetradon nigroviridis, genomic survey sequence. AL214492 AL214492.1 GI:7873311 GSS: genome survey sequence. SOURCE Tetradon nigroviridis ORGANISM Tetradon nigroviridis Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetradon. </p>
<p> 1 Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizmes,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W., and Weissenbach,J. Estimate of human gene number provided by genome-wide analysis using Tetradon nigroviridis DNA sequence Nat. Genet. 25 (2), 235-238 (2000) 20296633 10835645 </p>	<p> REFERENCE AUTHORS </p>
<p> TITLE JOURNAL MEDLINE PUBMED </p>	<p> TITLE JOURNAL MEDLINE PUBMED </p>
<p> Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C., Fizmes,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J. Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis Genome Res. 10 (7), 939-949 (2000) 20359837 10899143 3 (bases 1 to 1000) </p>	<p> REFERENCE </p>

Db 508 AAAAAAA 515

RESULT 15
AZ541311 845 bp DNA linear GSS 14-NOV-2000
LOCUS ENT0V65TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, genomic survey sequence.
ACCESSION AZ541311
VERSION AZ541311.1 GI:11148922
KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
REFERENCE 1 (bases 1 to 845)
AUTHORS Loftus, B., Van Aken, S. and Frazer, C.
TITLE Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: ML3-Reverse
Class: Shotgun
High quality sequence start: 42
High quality sequence stop: 612.
Location/Qualifiers
1..845

FEATURES

source
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: PHOS1; Site 1: Bat I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + l method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

ORIGIN

Query Match 5.5%; Score 40; DB 8; length 845;
Best Local Similarity 55.9%; Pred. No. 4.8;
Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 588 AGTTGTTTCAATCATATAAAGGAGAAATGTTGCTACATGTGACAGTTCTGGG 647
DB 407 AATTAAATTAATAATATCATTTTGAATTAATTTTATTTGTAATGACAGAACTGAA 466
QY 648 AGGCTTAACTTTTATGAGAGACTATCAATCATACAGATATGTCAAAAA 707
DB 467 AGTTCTTAATTTAATAGATGAATATATATCATATGATGCTTTTCTAACA 526
QY 708 AGACTAATTAATAACT 723
DB 527 TTATGAAAAAATAATAT 542

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 03:07:50 ; Search time 531 Seconds
(without alignments)
8060.213 Million cell updates/sec

Title: US-10-776-213-2
Perfect score: 723
Sequence: 1 cttcgattagcagcagcac.....aaaagactataatacat 723

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: N.GeneSeq_16Dec04:*
2: geneseqn1980s:*
3: geneseqn1990s:*
4: geneseqn2000s:*
5: geneseqn2001as:*
6: geneseqn2001bs:*
7: geneseqn2002as:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	723	100.0	723	4	AAD07475	Yeast pro
2	723	100.0	11427	4	AAD07497	YMR251AP
3	723	100.0	13073	4	AAD07493	YMR251AP
4	720.4	99.6	850	4	AAD07503	Yeast YMR
5	497.4	68.8	680	4	ABQ76446	S. cerevisiae
6	39	5.4	2000	8	ADAV1938	Rice gene
7	38.6	5.3	2311	6	ABSS1325	CDNA enco
8	38.4	5.3	544	11	ADT96256	CDNA enco
9	37.8	5.2	7736	5	AAS29224	Genomic s
10	37.8	5.2	7736	5	ABA16123	Human ner
11	37.8	5.2	7736	6	ABG68364	Human DNA
12	37.8	5.2	7736	10	ADG25486	Human CDN
13	37.4	5.2	386	5	ABV04394	Human pro
14	37	5.1	366	8	ABX58465	Human mus
15	37	5.1	366	8	ABX58465	CDNA enco
16	37	5.1	366	12	ADJ28192	Human mus
17	37	5.1	380	8	ABX36811	Bovine ES
18	37	5.1	400	6	ABQ59188	Human col
19	36.4	5.0	201	13	ADG41207	Human aut
20	36.4	5.0	591	13	ACN58226	Cotton gy

21	36.4	5.0	110300	13	ADS36499	Adg36499 Human aut
22	36.2	5.0	524	10	ABT22885	Abt22885 Breast ca
23	36	5.0	7624	6	ABJ34113	Abj34113 Human imm
24	35.8	5.0	553	13	ACN62578	ACN62578 Cotton de
25	35.8	5.0	815	2	AAZ00806	AAZ00806 Human sec
26	35.8	5.0	815	8	ADA39926	Ada39926 Human sec
27	35.8	5.0	815	8	ACC50510	Acc50510 Human sec
28	35.8	5.0	815	10	ADA56114	Ada56114 Gene enco
29	35.8	5.0	1143	8	ACA40062	ACA40062 Prokaryot
30	35.8	5.0	10517	13	ADS89438	Adg89438 Oligonuc
31	35.6	4.9	506	9	ACH17729	ACH17729 Human adu
32	35.6	4.9	5629	6	ABJ58962	Abj58962 Human tum
33	35.6	4.9	5988	6	ABK09744	Abk09744 Human ova
34	35.6	4.9	5988	10	ADH29006	Adh29006 Human chr
35	35.6	4.9	5988	12	ADH2483	Adh2483 Human chr
36	35.6	4.9	64796	13	ACN37231	ACN37231 Human per
37	35.4	4.9	844	5	ABV18166	ABV18166 Human pro
38	35.4	4.9	33353	4	AAK70003	AAK70003 Human imm
39	35.4	4.9	201239	8	ACA64924	ACA64924 Human PL2
40	35.2	4.9	281	6	ABV96856	ABV96856 Human pan
41	35.2	4.9	1837	2	AAZ41383	AAZ41383 Human nor
42	35.2	4.9	2000	8	ADA71938	Ada71938 Rice gene
43	35.2	4.9	2938	11	ACN88712	ACN88712 Breast ca
44	35.2	4.9	3197	13	ADS89390	Adg89390 Oligonuc
45	35.2	4.9	107330	12	ADQ97316	Adq97316 Mouse can

ALIGNMENTS

RESULT 1
ID AAD07475 standard; DNA; 723 BP.
AC AAD07475;
XX 10-AUG-2001 (first entry)
DT XX
XX Yeast promoter YMR251WA.
DE XX
XX Yeast; promoter; gene expression; fermentable carbon source; glucose;
KW non-fermentable carbon source; ethanol; yeast cell culture; ds.
XX Saccharomyces cerevisiae.
OS XX
PN MO200138549-A1.
XX
XX 31-MAY-2001.
XX
XX 17-NOV-2000; 2000WO-SE002277.
XX
XX 23-NOV-1999; 99SE-00004247.
XX
XX (ASTR) ASTRAZENECAB AB.
XX
XX Belfield G, Oakley C;
XX WPI, 2001-367697/38.
XX
XX New promoter sequences from Saccharomyces cerevisiae useful for
PT controlling expression of homologous and heterologous nucleic acid
PT expression in yeast cells.
XX
XX Claim 1; Page 67-68; 191pp; English.
XX
XX The invention relates to yeast promoters that are used to control the
XX expression of homologous and heterologous nucleic acids encoding proteins
XX and polypeptides in yeast cells. The yeast promoters are induced by a
XX fermentable carbon source such as glucose or a non-fermentable carbon
XX source such as ethanol or both. Therefore expression of nucleic acid
XX molecules encoding a polypeptide under the control of the novel yeast
XX promoters are regulated by varying the level of a fermentable carbon
XX source or a non-fermentable carbon source or both. The yeast promoters

CC are useful for, inter alia, the high level production of proteins or
 CC polypeptides in yeast cell culture. The present DNA sequence is
 CC Saccharomyces cerevisiae YMR251WA promoter related to the invention
 XX
 XX Sequence 723 BP; 199 A; 165 C; 166 G; 193 T; 0 U; 0 Other;

Query Match 100.0%; Score 723; DB 4; Length 723;
 Best Local Similarity 100.0%; Pred. No. 9.2e-211;
 Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CTTTGATTAGACGACACATCATAGATGCGGTCAATAAATACATACGAAAA 60
DB 1 CTTTGATTAGACGACACATCATAGATGCGGTCAATAAATACATACGAAAA 60
QY 61 ACCATAAAGAGCAACGATACCTACTTGGAAAGAAAGAGACCGCTTGAAGGGAGAT 120
DB 61 ACCATAAAGAGCAACGATACCTACTTGGAAAGAAAGAGACCGCTTGAAGGGAGAT 120
QY 121 GGGGGCTAAGAGTCACTTCTTTTCCCTTCCGCGGTCCGAGACCCGGACCCCTCCT 180
DB 121 GGGGGCTAAGAGTCACTTCTTTTCCCTTCCGCGGTCCGAGACCCGGACCCCTCCT 180
QY 181 CTCCTCCGACGATTTCTTCTTCAATATCTTCTTATCTTCCATCCGTTGAAGAAC 240
DB 181 CTCCTCCGACGATTTCTTCTTCAATATCTTCTTATCTTCCATCCGTTGAAGAAC 240
QY 241 GCACTATGATTAATGCTGCTGACATCTCAATGCTGATCTGTGATCTACAGT 300
DB 241 GCACTATGATTAATGCTGCTGACATCTCAATGCTGATCTGTGATCTACAGT 300
QY 301 GGTAAAGGACCGTGGCTCGGAAACGGTTCCTTCTGACAAATCTAGAACAGGGGCTACA 360
DB 301 GGTAAAGGACCGTGGCTCGGAAACGGTTCCTTCTGACAAATCTAGAACAGGGGCTACA 360
QY 361 GGTAAAGGACCGTGGCTCGGAAACGGTTCCTTCTGACAAATCTAGAACAGGGGCTACA 360
DB 361 GGTAAAGGACCGTGGCTCGGAAACGGTTCCTTCTGACAAATCTAGAACAGGGGCTACA 360
QY 421 ATAGGAGCGCGAGTTTATCGGCGAGCTCTACTTCTTCTTATTTGGTAAAGCCCTTTC 480
DB 421 ATAGGAGCGCGAGTTTATCGGCGAGCTCTACTTCTTCTTATTTGGTAAAGCCCTTTC 480
QY 481 TGTTCCTGCGCAGTGTGCTGACAGCTGCGCGGAGAACATATGATTAAGGAGTAAAC 540
DB 481 TGTTCCTGCGCAGTGTGCTGACAGCTGCGCGGAGAACATATGATTAAGGAGTAAAC 540
QY 541 TTTTCATGAGAGATTTAGCAAGCGGAAAAAACTATGCTAGCTGGGAGTTGTTTTTCA 600
DB 541 TTTTCATGAGAGATTTAGCAAGCGGAAAAAACTATGCTAGCTGGGAGTTGTTTTTCA 600
QY 601 TCATATTAAGAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGACGCTTAACTTT 660
DB 601 TCATATTAAGAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGACGCTTAACTTT 660
QY 661 TATTCGAGGAGACTATCAATCATACATATTTGTCAAAAAAAGAGACTAATATATA 720
DB 661 TATTCGAGGAGACTATCAATCATACATATTTGTCAAAAAAAGAGACTAATATATA 720
QY 721 CAT 723
DB 721 CAT 723

```

RESULT 2
 ID AAD07497 standard; DNA; 11427 BP.

XX AAD07497;
 XX 10-AUG-2001 (first entry)
 XX
 DE pYMR251AP DNA plasmid.
 XX

KW Yeast; promoter; gene expression; fermentable carbon source; glucose;
 KW non-fermentable carbon source; ethanol; yeast cell culture;
 KW pYMR251AP plasmid; ds.

XX Saccharomyces cerevisiae.
 OS Unidentified.
 OS Chimeric.

PN WO200138549-A1.

PD 31-MAY-2001.

PF 17-NOV-2000; 2000MO-SE002277.

PR 23-NOV-1999; 99SE-00004247.

PA (ASTR) ASTRAENCA AB.

PI Belfield G, Oakley C;

DR WPI, 2001-367697/38.

PT New promoter sequences from Saccharomyces cerevisiae useful for

PT controlling expression of homologous and heterologous nucleic acid

PS expression in yeast cells.

XX Example 4; Page 133-143; 191pp; English.

CC The invention relates to yeast promoters that are used to control the

CC expression of homologous and heterologous nucleic acids encoding proteins

CC and polypeptides in yeast cells. The yeast promoters are induced by a

CC fermentable carbon source such as glucose or a non-fermentable carbon

CC source such as ethanol or both. Therefore expression of nucleic acid

CC molecules encoding a polypeptide under the control of the novel yeast

CC promoters are regulated by varying the level of a fermentable carbon

CC source or a non-fermentable carbon source or both. The yeast promoters

CC are useful for, inter alia, the high level production of proteins or

CC polypeptides in yeast cell culture. The present sequence is pYMR251AP

CC plasmid related to the invention. This plasmid contains pPRB1 and yeast

XX YMR251WA promoter

Sequence 11427 BP; 3131 A; 2589 C; 2527 G; 3180 T; 0 U; 0 Other;

Query Match 100.0%; Score 723; DB 4; Length 11427;

Best Local Similarity 100.0%; Pred. No. 3.3e-210;

Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CTTTGATTAGACGACACATCATAGATGCGGTCAATAAATACATACGAAAA 60
DB 1 CTTTGATTAGACGACACATCATAGATGCGGTCAATAAATACATACGAAAA 60
QY 15 CTTTGATTAGACGACACATCATAGATGCGGTCAATAAATACATACGAAAA 74
DB 15 CTTTGATTAGACGACACATCATAGATGCGGTCAATAAATACATACGAAAA 74
QY 61 ACCATAAAGAGCAACGATACCTACTTGGAAAGAAAGAGACCGCTTGAAGGGAGAT 120
DB 61 ACCATAAAGAGCAACGATACCTACTTGGAAAGAAAGAGACCGCTTGAAGGGAGAT 120
QY 75 ACCATAAAGAGCAACGATACCTACTTGGAAAGAAAGAGACCGCTTGAAGGGAGAT 134
DB 75 ACCATAAAGAGCAACGATACCTACTTGGAAAGAAAGAGACCGCTTGAAGGGAGAT 134
QY 121 GGGGGCTAAGAGTCACTTCTTTTCCCTTCCGCGGTCCGAGACCGGAAACCCCTCT 180
DB 121 GGGGGCTAAGAGTCACTTCTTTTCCCTTCCGCGGTCCGAGACCGGAAACCCCTCT 180
QY 181 CTCCTCCGACGATTTCTTCTTCAATATCTTCTTATTTCTTATCCCGTTGAAGCAAC 240
DB 181 CTCCTCCGACGATTTCTTCTTCAATATCTTCTTATTTCTTATCCCGTTGAAGCAAC 240
QY 195 CTCCTCCGACGATTTCTTCTTCAATATCTTCTTATTTCTTATCCCGTTGAAGCAAC 254
DB 195 CTCCTCCGACGATTTCTTCTTCAATATCTTCTTATTTCTTATCCCGTTGAAGCAAC 254
QY 241 GCACTATGATTAATGCTGCTGACATCTCAATGCTGATCTGTGATCTACAGT 300
DB 241 GCACTATGATTAATGCTGCTGACATCTCAATGCTGATCTGTGATCTACAGT 300
QY 255 GCACTATGATTAATGCTGCTGACATCTCAATGCTGATCTGTGATCTACAGT 314
DB 255 GCACTATGATTAATGCTGCTGACATCTCAATGCTGATCTGTGATCTACAGT 314
QY 301 GGTAAAGGACCGTGGCTCGGAAACGGTTCCTTCTGACAAATCTAGAACAGGGGCTACA 360
DB 301 GGTAAAGGACCGTGGCTCGGAAACGGTTCCTTCTGACAAATCTAGAACAGGGGCTACA 360
QY 315 GGTAAAGGACCGTGGCTCGGAAACGGTTCCTTCTGACAAATCTAGAACAGGGGCTACA 374
DB 315 GGTAAAGGACCGTGGCTCGGAAACGGTTCCTTCTGACAAATCTAGAACAGGGGCTACA 374
QY 361 GTCTCGATTAATGAATTAATAGCGCATTTTGTCTAGCGCCGCGGCGCCGCTTTCCCA 420
DB 361 GTCTCGATTAATGAATTAATAGCGCATTTTGTCTAGCGCCGCGGCGCCGCTTTCCCA 420

```

Db 375 GTCTCGATATAGAAATTAAGCCATTTTGTAGCCGCCGCCGCTTTCCCA 434
QY 421 ATAGGAGGCGGAGTTATATGCGGAGCTTACTTCTTATTTGGGTAAGCCCTTTC 480
Db 435 ATAGGAGGCGGAGTTATATGCGGAGCTTACTTCTTATTTGGGTAAGCCCTTTC 494
QY 481 TGTTCGCGCAGTGTGCTGCAAGCTGCGCGAGAACATATGATTAAGGATTAAC 540
Db 495 TGTTCGCGCAGTGTGCTGCAAGCTGCGCGAGAACATATGATTAAGGATTAAC 554
QY 541 TTTTCGATGAGAAATTAAGCAAGCGAAAAAACTATGCTGAGTGTGTTTCAA 600
Db 555 TTTTCGATGAGAAATTAAGCAAGCGAAAAAACTATGCTGAGTGTGTTTCAA 614
QY 601 TCATATTAAGGAGAAATTTGCTCCTATGTCAGTTTCTGGAGCTTAACCTT 660
Db 615 TCATATTAAGGAGAAATTTGCTCCTATGTCAGTTTCTGGAGCTTAACCTT 674
QY 661 TATTGAGAGACTATCAATATCAATACAGATATTTGTCAAAAAAAAGACTAATAA 720
Db 675 TATTGAGAGACTATCAATATCAATACAGATATTTGTCAAAAAAAAGACTAATAA 734
QY 721 CAT 723
Db 735 CAT 737

RESULT 3

AAD07493 standard; DNA; 13073 BP.

AAD07493;

10-AUG-2001 (first entry)

pYMR251AP+luc sequence DNA construct.

Yeast; promoter; gene expression; fermentable carbon source; glucose;

non-fermentable carbon source; ethanol; yeast cell culture;

pYMR251AP plasmid; luciferase gene; ds.

Saccharomyces cerevisiae.

Unidentified.

Chimeric.

MO200138549-A1.

31-MAY-2001.

17-NOV-2000; 2000MO-SR002277.

23-NOV-1999; 99SE-00004247.

(ASTR) ASTRAZENECA AB.

Belfield G, Oakley C;

WPI, 2001-367697/38.

New promoter sequences from Saccharomyces cerevisiae useful for

controlling expression of homologous and heterologous nucleic acid

expression in yeast cells.

Example 3; Page 88-99; 191pp; English.

The invention relates to yeast promoters that are used to control the

expression of homologous and heterologous nucleic acids encoding proteins

and polypeptides in yeast cells. The yeast promoters are induced by a

fermentable carbon source such as glucose or a non-fermentable carbon

source such as ethanol or both. Therefore expression of nucleic acid

molecules encoding a polypeptide under the control of the novel yeast

promoters are regulated by varying the level of a fermentable carbon

source or a non-fermentable carbon source or both. The yeast promoters

CC are useful for, inter alia, the high level production of proteins or
CC polypeptides in yeast cell culture. The present sequence is pYMR251AP+luc
CC sequence DNA construct related to the invention. The pYMR251AP+luc
CC sequence contains pPRB1 plasmid, luc backbone (luciferase gene) and yeast
CC YMR251WA promoter
XX

SQ Sequence 13073 BP; 3584 A; 2949 C; 2934 G; 3606 T; 0 U; 0 Other;

Query Match 100.0%; Score 723; DB 4; Length 13073;

Best Local Similarity 100.0%; Pred. No. 3.5e-210; Mismatches 0; Gaps 0;

Matches 723; Conservative 0; Indels 0; Gaps 0;

QY 1 CTTTCGATTAAGACGACACATACATAGACGTCGTATTAATAATACATACGAAAA 60
Db 16 CTTTCGATTAAGACGACACATACATAGACGTCGTATTAATAATACATACGAAAA 75
QY 61 ACCATTAAGAGCAAGCGATACCTTCTGGAAGAAAAGAGACGCTTTGAAGGGGAT 120
Db 76 ACCATTAAGAGCAAGCGATACCTTCTGGAAGAAAAGAGACGCTTTGAAGGGGAT 135
QY 121 GGGGGCTAAGAGTCACTTCTTTCCTTCCGCGGCTCGGACCGGGACCCCTCCT 180
Db 136 GGGGGCTAAGAGTCACTTCTTTCCTTCCGCGGCTCGGACCGGGACCCCTCCT 195
QY 181 CTCGCCGACAGATTTCTTCTTCAATCTTCTTATTCCTATCCGTTGAAGAACCC 240
Db 196 CTCGCCGACAGATTTCTTCTTCAATCTTCTTATTCCTATCCGTTGAAGAACCC 255
QY 241 GCACTATGACTAATATGCTGCAATCTTCATATGCTGCAATCTTCATCTCAGCT 300
Db 256 GCACTATGACTAATATGCTGCAATCTTCATATGCTGCAATCTTCATCTCAGCT 315
QY 301 GGTAAACGAGACCGTGGCTCGAAGACGTTCTTCGTAATTTCTAGAACAGGGGCTACA 360
Db 316 GGTAAACGAGACCGTGGCTCGAAGACGTTCTTCGTAATTTCTAGAACAGGGGCTACA 375
QY 361 GTCTCGATTAATAGAAATTAATAGCGCATTTTCTAGCGCGCGCGCCGCTTCCCA 420
Db 376 GTCTCGATTAATAGAAATTAATAGCGCATTTTCTAGCGCGCGCGCCGCTTCCCA 435
QY 421 ATAGGAGGCGGAGTTATATGCGGAGCTTACTTCTTCTTATTTGGGTAAGCCCTTTC 480
Db 436 ATAGGAGGCGGAGTTATATGCGGAGCTTACTTCTTCTTATTTGGGTAAGCCCTTTC 495
QY 481 TGTTCGCGCAGTGTGCTGCAAGCTGCGCGAGAACATATGATTAAGGATTAAC 540
Db 496 TGTTCGCGCAGTGTGCTGCAAGCTGCGCGAGAACATATGATTAAGGATTAAC 555
QY 541 TTTTCGATGAGAAATTTAGCAAGCGAAAAAACTATGCTAGCTGGAGTTGTTTCAA 600
Db 556 TTTTCGATGAGAAATTTAGCAAGCGAAAAAACTATGCTAGCTGGAGTTGTTTCAA 615
QY 601 TCATATTAAGGAGAAATTTGCTCACTATGTCAGTTTCTGGAGCTTTAACTTT 660
Db 616 TCATATTAAGGAGAAATTTGCTCACTATGTCAGTTTCTGGAGCTTTAACTTT 675
QY 661 TATTGAGAGACTATCAATATCAATACAGATATTTGTCAAAAAAAAGACTAATAA 720
Db 676 TATTGAGAGACTATCAATATCAATACAGATATTTGTCAAAAAAAAGACTAATAA 735
QY 721 CAT 723
Db 736 CAT 738

RESULT 4

AAD07503 standard; DNA; 850 BP.

AAD07503;

10-AUG-2001 (first entry)

XX

```
DE Yeast YMR251WA promoter region.
XX
XX Yeast; promoter; gene expression; fermentable carbon source; glucose;
KM non-fermentable carbon source; ethanol; yeast cell culture; ds.
XX
XX Saccharomyces cerevisiae.
XX
FH Key Location/Qualifiers
FT 1..64
FT /tag= a
FT /product= "Yeast YMR251W open reading frame encoding
FT polypeptide"
FT complement(78..95)
FT /tag= c
FT /bound moiety= "YMR251WA forward PCR primer"
FT /partial
FT 786..812
FT /tag= d
FT /bound moiety= "YMR251WA reverse PCR primer"
FT /partial
CDS
FT 801..850
FT /tag= b
FT /product= "Yeast YMR251WA open reading frame encoding
FT polypeptide"
XX
XX WO200138549-A1.
XX
XX 31-MAY-2001.
XX
XX 17-NOV-2000; 2000WO-SE002277.
XX
XX 23-NOV-1999; 99SE-00004247.
XX
XX (ASTR ) ASTRAZENECA AB.
XX
XX Belfield G, Oakley C;
XX
XX WPI; 2001-367697/38.
XX
XX
XX New promoter sequences from Saccharomyces cerevisiae useful for
XX controlling expression of homologous and heterologous nucleic acid
XX expression in yeast cells.
XX
XX Example 3; Fig 14; 191pp; English.
XX
XX
XX The invention relates to yeast promoters that are used to control the
XX expression of homologous and heterologous nucleic acids encoding proteins
XX and polypeptides in yeast cells. The yeast promoters are induced by a
XX fermentable carbon source such as glucose or a non-fermentable carbon
XX source such as ethanol or both. Therefore expression of nucleic acid
XX molecules encoding a polypeptide under the control of the novel yeast
XX promoters are regulated by varying the level of a fermentable carbon
XX source or a non-fermentable carbon source or both. The yeast promoters
XX are useful for, inter alia, the high level production of proteins or
XX polypeptides in yeast cell culture. The present DNA sequence is
XX Saccharomyces cerevisiae YMR251WA promoter region related to the
XX invention
XX
XX Sequence 850 BP; 225 A; 199 C; 193 G; 233 T; 0 U; 0 Other;
SQ
XX
XX Query Match 99.6%; Score 720.4; DB 4; Length 850;
XX Best Local Similarity 99.9%; Pred. No. 6.2e-210;
XX Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 CTTTGATTAGGACGACACATCATGATGCGGTATATAAATACACTACGGAATA 60
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 78 CTTTCGATTAGCAGCAGACATCATGATGCGGTATATAAATACACTACGGAATA 137
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 61 ACCATTAAGAGCAAGCGATACCTACTTGGAAGAAAGAGACGCTTGAAGGGGAT 120
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX Db 138 ACCATTAAGAGCAAGCGATACCTACTTGGAAGAAAGAGACGCTTGAAGGGGAT 197
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 121 GGGGGCTAAGAGTCACTTCTTTTCCTTCGGGCTCGGACCGGACCCCTCT 180
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 198 GGGGGCTAAGAGTCACTTCTTTCCCTTCGCGGCTCGGACCGGACCCCTCT 257
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 181 CTCGCCGACGATTTCTTCCTTTCATATCTTCTTTATTCCTATCCCGTTGAACAAC 240
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 258 CTCGCCGACGATTTCTTCCTTTCATATCTTCTTTATTCCTATCCCGTTGAACAAC 317
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 241 GCACATGACTAATAATGCTGCTGACATCTGCATGCTGATCTTGATCTTCACAGT 300
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 318 GCACATGACTAATAATGCTGCTGACATCTGCATGCTGATCTTGATCTTCACAGT 377
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 301 GGTAAAGGACCGCTGCTGCGAAAGGTTCTTCTGTAACAATTCTGAACAGGGCTACA 360
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 378 GGTAAAGGACCGCTGCTGCGAAAGGTTCTTCTGTAACAATTCTGAACAGGGCTACA 437
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 361 GTCCTGATATAGAAATAAAGCGCATTTTGTGACGCGCGCGCGCGCGGTTTCCCA 420
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 438 GTCTCGATATATAGAAATAAAGCGCATTTTGTGACGCGCGCGCGCGGTTTCCCA 497
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 421 ATAGGAGCGCGAGTTTATCGCGGAGCTTAATTCTTCTATTTGGTAAAGCCCTTTC 480
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 498 ATAGGAGCGCGAGTTTATCGCGGAGCTTAATTCTTCTATTTGGTAAAGCCCTTTC 557
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 481 TGTTTTGGCCAGTGTGCTGTCAGAGCTGCGCGGAAACATAGATTAAGGATGTAC 540
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 558 TGTTTTGGCCAGTGTGCTGTCAGAGCTGCGCGGAAACATAGATTAAGGATGTAC 617
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 541 TTTGATGAGAAATTTAGCAAGCGGAAAAAAACTATGCTGCTGAGGTTGTTTTCAA 600
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 618 TTTGATGAGAAATTTAGCAAGCGGAAAAAAACTATGCTGCTGAGGTTGTTTTCAA 677
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 601 TCATATTAAGGAGAAATTTGTGCTCATATGTGACAGTTTCTGGAAGCTTTAACTTT 660
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 678 TCATATTAAGGAGAAATTTGTGCTCATATGTGACAGTTTCTGGAAGCTTTAACTTT 737
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 661 TATTGCAAGAGACTTCAATCATACAGATATTGTCAAAAAAAGACTAATTAATTA 720
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 738 TATTGCAAGAGACTTCAATCATACAGATATTGTCAAAAAAAGACTAATTAATTA 797
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 721 CA 722
XX ||
Db 798 AA 799
XX ||
XX
XX RESULT 5
XX ID ABQ76446 standard; cDNA; 680 BP.
XX
XX ABQ76446;
XX
XX 21-NOV-2002 (first entry)
XX
XX S. cerevisiae BAX-associated cDNA fragment SEQ ID 317.
XX
XX Bax; Bax-resistance; cyrostatic; fungicide; immunosuppressive; virucide;
XX vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
XX apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
XX neurodegeneration; cell death; ss.
XX
XX Saccharomyces cerevisiae.
XX
XX WO200264766-A2.
XX
XX 22-AUG-2002.
XX
XX 21-DEC-2001; 2001WO-EP015398.
XX
XX 22-DEC-2000; 2000EP-00870318.
XX
XX 04-JAN-2001; 2001EP-00870002.
XX
XX 09-JAN-2001; 2001EP-00870003.
XX
XX (JANC ) JANSSEN PHARM NV.
```

PI Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX MPI: 2002-667002/71.
DR P-PSDB; ABG93180.
XX
XX New isolated nucleic acid representing a synthetic BAX-gene, useful as
PT medicament for treating, preventing and/or alleviating yeast or fungal
PT infections or proliferative disorders, or for preventing apoptosis in
PT certain diseases.
XX
XX Claim 36, fig 1; 344p; English.
XX
XX This invention describes a novel nucleic acid representing a synthetic
CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
CC resistant yeast or fungi, identifying, or obtaining and identifying
CC Candida spp. sequences that are differentially expressed in a pathway
CC eventually leading to programmed cell death or identifying inhibitors or
CC inhibitor sequences of Bax-induced cell death. The products of the
CC invention have cytostatic, fungicide, immunosuppressive, virocidic and
CC vasotropic activity and can be used in vaccines or for gene therapy. The
CC isolated nucleic acid, polypeptides, pharmaceutical compositions,
CC antiense molecules and antibodies are useful as medicaments or in
CC preparing a medicament for treating, preventing and/or alleviating
CC diseases associated with yeast or fungi or proliferative disorders, such
CC as cancer, or for preventing apoptosis in certain diseases. The compounds
CC or polypeptides, or the genetically modified organism are useful for
CC preparing a medicament for modifying the endogenic flora of humans and
CC other mammals. The vaccine is useful for immunising against yeast or
CC fungal infections. Apoptosis-related diseases include autoimmune disease,
CC leishmania, diseases related with viral infections or neurodegenerations.
CC This sequence represents a polynucleotide associated with the Bax gene
CC described in the disclosure of the invention
XX

Sequence 680 BP; 180 A; 142 C; 166 G; 192 T; 0 U; 0 Other;

Query Match 68.8%; Score 497.4; DB 6; Length 680;
Best Local Similarity 99.8%; Pred. No. 1.3e-141;
Matches 498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 224 ATCCCGTTGAAGACCGCACTATGACTTAATGTGCTGACACTCTCCATGCTGACT 283
DB 1 ATCCCGTTGAAGACCGCACTATGACTTAATGTGCTGACACTCTCCATGCTGACT 60
QY 284 TGTGTATCTCACTGAGTGAAGGACCGTGGCTGGAAAGGCTTCTTGGACAAAT 343
DB 61 TGTGTATCTCACTGAGTGAAGGACCGTGGCTGGAAAGGCTTCTTGGACAAAT 120
QY 344 CTAGAACAGGGGCTACAGTCTCGATATATGAATATAGGCAATTTTGGTACGGCGCC 403
DB 121 CTAGAACAGGGGCTACAGTCTCGATATATGAATATAGGCAATTTTGGTACGGCGCC 180
QY 404 GCGGCGCCGCTTCCCAATGAGGAGCGCAATTATCGCGGAGCTTACTTCTCTAT 463
DB 181 GCGGCGCCGCTTCCCAATGAGGAGCGCAATTATCGCGGAGCTTACTTCTCTAT 240
QY 464 TTGGGTAAGCCCTTCTGTTTCCGCACTGCTGAGGCTGCGCGGAAACATA 523
DB 241 TTGGGTAAGCCCTTCTGTTTCCGCACTGCTGAGGCTGCGCGGAAACATA 300
QY 524 GTGATTAAGGATGTAATCTTCGATGAGAAATAGCAAGGAAAAAACTATGCTAGC 583
DB 301 GTGATTAAGGATGTAATCTTCGATGAGAAATAGCAAGGAAAAAACTATGCTAGC 360
QY 584 TGGGAGTTGTTTTCATCATATATAAAAGGAAATTTGTTCTCACTATGTGCAAGTTTC 643
DB 361 TGGGAGTTGTTTTCATCATATATAAAAGGAAATTTGTTCTCACTATGTGCAAGTTTC 420
QY 644 TGGGAGCTTAACTTTATTTGAGAGGAACTATCAATCATACAGATTTGTCAAAAAA 703
DB 421 TGGGAGCTTAACTTTATTTGAGAGGAACTATCAATCATACAGATTTGTCAAAAAA 480
QY 704 AAAAGACTAATATACA 722
|||||

DB 481 AAAAGACTAATATATAAA 499

RESULT 6
ADA71938
ID ADA71938 standard; DNA; 2000 BP.

XX ADA71938;

XX 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;
KM gene; de.

XX Oryza sativa.

XX WO200300898-A1.

XX 03-JUN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;

XX MPI: 2003-175290/17.

PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

PS Claim 27; SEQ ID NO 5263; 899p; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 5.4%; Score 39; DB 8; Length 2000;
Best Local Similarity 8.1%; Pred. No. 0.66;
Matches 45; Conservative 258; Mismatches 248; Indels 2; Gaps 1;

QY 75 AGCATACCTACTTGGAGAAAGAGACCGCTTGAAGGGAGTGGGGCTAAGAACT 134
DB 184 AATKSGSMKRMWMSCGSGGGRSAVSRYGTSTKTYKMYTYSASRCRAYMTTS 243
QY 135 CATTCACCTTCTTTCCCTCGCGGTCGGACCGGAGCCCTCTCCCGACAGATT 194
DB 244 YSMAGSSTWCSKRSMMKMKRMSRSTGWSMSTKMMCTVYKSYSRMCTMR 303
QY 195 TCTTCCTTCATATCTCTCTTATTCCTATCCGTTGAAGCAACGCACTATGACTAAA 254
DB 304 GGGWGTATYTWGRGVMSRAAMMYKKMYVRYGKMKRGWAGRMMSRCWMSKACYMR 363
QY 255 TGTGTGACATCTCCATGCGCTGAGCTTGTGTATATCAAGTGGTAAGGACCGCT 314
DB 364 WMMWMTTRRRWAKKSRTSRKSKKMKRKYKRMGRYSRMSCKRARMMKCRSGRA 423
:::|||||

QY 315 GGCTCGAAGCGGCTTCCTTCTGTGACATTTGTAGAACAGGGGCTACAGTCTCGATTAATAGA 374
 Db 424 WKAGCGGCMTCRMKSKYGMWRMKSKMKNASKXKMMSSRMVPRKKCSRTTMMGKTRGMWG 483
 QY 375 ATATATAGGCGCATTTTTCCTAGAGCGCGCGCGCGCGCGCTTCCCATATAGGAGGGCGCA- 433
 Db 484 TMRGCRFKGSKMKRCKRRRRMRMRMRKRYMSARVYMRCAARKKSYSAAARACW 543
 QY 434 -GTTTATCGGCGAGGCTCTACTTCTTCTTATTTGGTAAGCCCGCTTTGTTTCCGCCA 492
 Db 544 YRKGYVMGMWMMKRYKRYMYMYMMWMMYKXKSKCMYCMSSYYACMSARAAKMK 603
 QY 493 GTGCTTCTGCGAGGCTGCGCCGGAACAATAGTATAGGATGATTACTTTCATAGAG 552
 Db 604 RSRMSAMSKMSRSGSRCKRCKASKRSSAKRYMMGWTGSGSRMSRMSYTCYMRKMGSMKS 663
 QY 553 AATTACAGCGGAGAAAAAATCTATGCTGCTGCGAGTGTCTTTCATCATATAAAG 612
 Db 664 TCTWMTYMSKYTYAKGYWRYRYRAMCMYMMWYRYRYSYMTYMAWYTSSTRMAMTGM 723
 QY 613 GAGAAATTTGTC 625
 Db 724 KYSGRYTSMYKY 736

RESULT 7
 ID ABS51325 standard; cDNA; 2311 BP.

XX ABS51325;

XX 21-OCT-2002 (first entry)

XX cDNA encoding human secretory protein #23.

XX Human; secretory polypeptide; SPTM; actinic keratosis; arteriosclerosis;
 XX bursitis; cirrhosis; hepatitis; polycythaemia vera; anaemia; psoriasis;
 XX primary thrombocytopenia; cancer; adenocarcinoma; leukemia; myeloma;
 XX sarcoma; immune system disorder; acquired immunodeficiency syndrome;
 XX AIDS; allergy; asthma; Crohn's disease; diabetes mellitus; gout;
 XX glomerulonephritis; Goodpasture's syndrome; thyroiditis; pancreatitis;
 XX hepatitis; multiple sclerosis; osteoporosis; Reiter's syndrome;
 XX rheumatoid arthritis; neurological disorder; epilepsy; stroke; dementia;
 XX Alzheimer's disease; Pick's disease; Huntington's disease; mood; anxiety;
 XX Parkinson's disease; central nervous system disorder; mental disorder;
 XX schizophrenic disorder; amnesia; Tourette's disorder; transgenic animal;
 XX gene therapy; gene; ss.

OS Homo sapiens.

XX WO200257304-A2.

XX 25-JUL-2002.

XX 15-JAN-2002; 2002WO-US001340.

XX 16-JAN-2001; 2001US-0261864P.
 XX 16-JAN-2001; 2001US-0261865P.
 XX 16-JAN-2001; 2001US-0261979P.
 XX 16-JAN-2001; 2001US-0261981P.
 XX 17-JAN-2001; 2001US-0262164P.
 XX 17-JAN-2001; 2001US-0262208P.
 XX 17-JAN-2001; 2001US-0263131P.
 XX 19-JAN-2001; 2001US-0262599P.
 XX 19-JAN-2001; 2001US-0262760P.
 XX 19-JAN-2001; 2001US-0263063P.
 XX 19-JAN-2001; 2001US-0263066P.
 XX 19-JAN-2001; 2001US-0263069P.
 XX 19-JAN-2001; 2001US-0263070P.
 XX 19-JAN-2001; 2001US-0263074P.
 XX 19-JAN-2001; 2001US-0263076P.
 XX 19-JAN-2001; 2001US-0263077P.
 XX 19-JAN-2001; 2001US-0263129P.

XX (INCY-) INCYTE GENOMICS INC.
 PA Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL;
 PI Dam TC, Liu TP, Harris B, Flores V, Daffo A, Marwaha R, Chen AJ;
 PI Chang SC, Gerstlin EH, Peralta CH, David WH, Lewis SA;
 XX WPI; 2002-590716/63.
 XX P-PSDB; ABG69833.

PT New purified secretory polypeptides and polynucleotides, useful in the
 PT diagnosis, study, prevention or treatment of diseases associated with
 PT decreased expression of functional secretory molecules, e.g. AIDS, cancer
 PT or allergies.

PS Claim 1, Page 260-261; 340pp; English.

XX The invention describes an isolated polynucleotide a naturally occurring
 CC polynucleotide sequence at least 90 % identical to it, a polynucleotide
 CC complementary to it or an RNA equivalent of it. The purified secretory
 CC polypeptides (SPTM) and polynucleotides are useful in the diagnosis,
 CC study, prevention or treatment of diseases associated with decreased
 CC expression of functional SPTM, e.g. actinic keratosis, arteriosclerosis,
 CC bursitis, cirrhosis, hepatitis, polycythaemia vera, primary
 CC thrombocytopenia, anaemia, psoriasis, cancers including adenocarcinoma,
 CC leukemia, myeloma or sarcoma, immune system disorder such as acquired
 CC immunodeficiency syndrome (AIDS), allergies, asthma, Crohn's disease,
 CC diabetes mellitus, glomerulonephritis, Goodpasture's syndrome, gout,
 CC Hashimoto's thyroiditis, hepatitis, multiple sclerosis, osteoporosis,
 CC pancreatitis, Reiter's syndrome, autoimmune thyroiditis or rheumatoid
 CC arthritis, neurological disorders such as epilepsy, stroke, Alzheimer's
 CC disease, Pick's disease, Huntington's disease, dementia, Parkinson's
 CC disease, other developmental disorder of the central nervous system,
 CC mental disorder including mood, anxiety or schizophrenic disorder,
 CC amnesia or Tourette's disorder. The polynucleotides may be used in
 CC hybridisation and amplification technologies, e.g. in assessing gene
 CC expression patterns, to develop a transcript image for a particular cell
 CC or tissue, or to create transgenic animals to model human disease. This
 CC sequence encodes a human secretory protein isolated in the invention
 XX

SQ Sequence 2311 BP, 824 A, 325 C; 473 G; 689 T; 0 U; 0 Other;

Query Match 5.3%; Score 38.6; DB 6; Length 2311;
 Best Local Similarity 53.7%; Pred. No. 0.94;
 Matches 80; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 551 AGAATTACGACGCGGAAAAAACAATAGCTAGCTGCGAGTGTCTTTCATCATATAA 610
 Db 345 AGTTTAAACAAATTCCTTAAATAATTAATTAATTAATTAATTAATTAATTA 286
 QY 611 GGGAGAAATGTGTCACATAATGACAGTTTCTGGAGCTCTTAATTTATTGACAG 670
 Db 285 CAGCTTAGTTTCTTACTTTTAAATTAATTTGGGTGATGTTTAATTAATTAATTTGACAT 226
 QY 671 GACTATCAAAATCATACAGATATGTCAAA 699
 Db 225 TAATCATTTCTCAGTGTATATATTTCCAA 197

RESULT 8
 ID ADT96256/c
 XX ADT96256 standard; cDNA; 544 BP.

XX ADT96256;

XX 16-DEC-2004 (first entry)

XX Colon cancer associated human cDNA sequence #1763.

XX Colón cancer; T cell; tumour protein; C634S; C637S; C640S; C636S;
 XX humoral immune response; cellular immune response; cytotoxic;
 XX immunostimulant; human; ss.

OS	Homo sapiens.
XX	
PN	US2003087818-A1.
PD	
XX	08-MAY-2003.
XX	
PF	01-FEB-2002; 2002US-00066543.
XX	
PR	02-FEB-2001; 2001US-0267400P.
PR	07-FEB-2001; 2001US-0267382P.
PR	11-MAY-2001; 2001US-0290323P.
PR	12-JUL-2001; 2001US-0305265P.
PR	16-AUG-2001; 2001US-0313077P.
XX	
PA	(CORI-) CORIXA CORP.
PI	
PI	Jiang Y, Chenault RA, Xu J, Indrias CY, Iodes MU, Secrist H;
PI	Catter D, Fanger GR, Smith CL, Durham M, Stolk JA;
DR	WPI; 2003-040540/03.
XX	
PT	New isolated nucleic acids and polypeptides capable of eliciting a
PT	humoral and/or cellular immune response, useful for diagnosing,
PT	preventing or treating cancer, particularly colon cancer.
XX	
PS	Claim 1; SEQ ID NO 1775; 87pp; English.
XX	
CC	The invention relates to polynucleotide and polypeptide sequences
CC	associated with cancer, particularly colon cancer. Also disclosed are (i)
CC	an expression vector comprising the polynucleotide, (ii) a host cell
CC	transformed or transfected with the expression vector, (iii) an isolated
CC	antibody, or its antigen-binding fragment, which specifically binds to
CC	the polypeptide, (iv) a method of detecting or determining the presence
CC	of cancer in a patient, (v) a fusion protein comprising at least one of
CC	the polypeptides, (vi) an oligonucleotide that hybridises to the
CC	polynucleotide sequence under highly stringent conditions, and (vii) a
CC	method of stimulating and/or expanding T cells specific for a tumour
CC	protein. The polypeptide specifically comprises the amino acid sequence
CC	of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC	frames (ORFs) of C636S. These polypeptides are encoded by the
CC	polynucleotide sequences, where both are capable of eliciting a humoral
CC	and/or cellular immune response. The polynucleotides, polypeptides, and
CC	antibodies are useful for diagnosing, preventing or treating cancer,
CC	particularly colon cancer. The polynucleotide and polypeptide sequences
CC	are also useful in DNA strand invasion, antisense inhibition, mutational
CC	analysis, nucleic acid purification, isolation of transcriptionally
CC	active genes, blocking or transcription factor binding, genome cleavage,
CC	or in situ hybridisation, and as enhancers of transcription or
CC	biomarkers. This sequence represents a human colon cancer associated
CC	cDNA. Note: The sequence data for this patent was obtained in electronic
CC	format directly from the USPTO web site at seqdata.uspto.gov
XX	
XX	Sequence 544 BP; 181 A; 70 C; 73 G; 216 T; 0 U; 4 Other;
XX	
QY	Query Match 5.3%; Score 38.4; DB 11; Length 544;
QY	Best Local Similarity 57.5%; Pred. No. 0.55;
QY	Matches 69; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
Db	
QY	601 TCATATTAAGAGGAGAAATTGTTGCTCACTATGACAGTTTGGAGCGTCTTAATTT 660
Db	
QY	126 TCATTGGAATGTGGAATTCAAATACAGCGTATGTAAATTTTTCATATGTCATATTTT 67
Db	
QY	661 TATTGCGAGGAGACTATCAATCATCATATTTTGTCAAAAAAAAAAGACTAATAATTA 720
Db	
QY	66 GAAAAAATTAATTTTAAATAATCATTCATCAAAAAAATAAAAAAAAAAAAAAAAAA 7
Db	
XX	
XX	RESULT 9
XX	AAS29224/C
ID	AAS29224 standard; DNA; 7736 BP.
XX	
XX	AAS29224;
XX	

DT	21-NOV-2001	(first entry)
XX		
DE	Genomic sequence #67 encoding novel human DNA-binding protein.	
XX		
KW	Human; DNA-binding protein; histone; chromo domain protein;	
KW	chromatin organisation modifier; y-box binding protein; DNA organisation;	
KW	gene transcription; malignant disease; autoimmune disorder;	
KW	rheumatic diseases; genetic abnormality; infectious disease;	
KW	neurological disorder; gene therapy; immunomodulatory; anti-HIV;	
KW	anti rheumatic; anti microbial; cytostatic; ds.	
OS	Homo sapiens.	
XX		
PN	WO2001S516Z-A1.	
XX		
PD	02-AUG-2001.	
PF	17-JAN-2001; 2001WO-US001305.	
XX		
PR	31-JAN-2000; 2000US-0179065P.	
PR	04-FEB-2000; 2000US-0180628P.	
PR	24-FEB-2000; 2000US-0184664P.	
PR	02-MAR-2000; 2000US-0186350P.	
PR	16-MAR-2000; 2000US-0189874P.	
PR	17-MAR-2000; 2000US-0190076P.	
PR	18-APR-2000; 2000US-0198123P.	
PR	19-MAY-2000; 2000US-0205515P.	
PR	07-JUN-2000; 2000US-0209467P.	
PR	28-JUN-2000; 2000US-0214886P.	
PR	30-JUN-2000; 2000US-0215135P.	
PR	07-JUL-2000; 2000US-0216647P.	
PR	11-JUL-2000; 2000US-0216880P.	
PR	11-JUL-2000; 2000US-0217496P.	
PR	14-JUL-2000; 2000US-0218290P.	
PR	26-JUL-2000; 2000US-0220963P.	
PR	26-JUL-2000; 2000US-0220964P.	
PR	14-AUG-2000; 2000US-0224518P.	
PR	14-AUG-2000; 2000US-0224519P.	
PR	14-AUG-2000; 2000US-0225213P.	
PR	14-AUG-2000; 2000US-0225214P.	
PR	14-AUG-2000; 2000US-0225266P.	
PR	14-AUG-2000; 2000US-0225267P.	
PR	14-AUG-2000; 2000US-0225268P.	
PR	14-AUG-2000; 2000US-0225270P.	
PR	14-AUG-2000; 2000US-0225447P.	
PR	14-AUG-2000; 2000US-0225757P.	
PR	14-AUG-2000; 2000US-0225758P.	
PR	14-AUG-2000; 2000US-0225759P.	
PR	18-AUG-2000; 2000US-0226279P.	
PR	22-AUG-2000; 2000US-0226681P.	
PR	22-AUG-2000; 2000US-0226868P.	
PR	22-AUG-2000; 2000US-0227182P.	
PR	23-AUG-2000; 2000US-0228924P.	
PR	01-SEP-2000; 2000US-0228287P.	
PR	01-SEP-2000; 2000US-0228343P.	
PR	01-SEP-2000; 2000US-0228344P.	
PR	01-SEP-2000; 2000US-0229345P.	
PR	05-SEP-2000; 2000US-0229509P.	
PR	05-SEP-2000; 2000US-0229513P.	
PR	06-SEP-2000; 2000US-0230437P.	
PR	06-SEP-2000; 2000US-0230438P.	
PR	08-SEP-2000; 2000US-0231242P.	
PR	08-SEP-2000; 2000US-0231243P.	
PR	08-SEP-2000; 2000US-0231244P.	
PR	08-SEP-2000; 2000US-0231413P.	
PR	08-SEP-2000; 2000US-0231414P.	
PR	08-SEP-2000; 2000US-0232080P.	
PR	08-SEP-2000; 2000US-0232081P.	
PR	12-SEP-2000; 2000US-0231968P.	
PR	14-SEP-2000; 2000US-0232397P.	
PR	14-SEP-2000; 2000US-0232398P.	

KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

OS Homo sapiens.

PN WO200159063-A2.

XX 16-AUG-2001.

PF 17-JAN-2001; 2001WO-US001334.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214866P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 18-AUG-2000; 2000US-0225759P.

PR 22-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234979P.

PR 25-SEP-2000; 2000US-0234988P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235634P.

PR 27-SEP-2000; 2000US-0235636P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.

PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.

PR 17-NOV-2000; 2000US-0249246P.

PR 17-NOV-2000; 2000US-0249257P.

PR 17-NOV-2000; 2000US-0249297P.

PR 17-NOV-2000; 2000US-0249299P.

PR 17-NOV-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250391P.

PR 01-DEC-2000; 2000US-0251160P.

PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.

PR 06-DEC-2000; 2000US-0251479P.

PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.

PR 08-DEC-2000; 2000US-0251989P.

PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-541565/60.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
PS Disclosure; SEQ ID NO 8454; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABA14678-ABA18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/pub/publ/published_pct_sequences
XX
SQ Sequence 7736 BP; 1672 A; 1981 C; 2164 G; 1919 T; 0 U; 0 Other;

Query Match 5.2%; Score 37.8; DB 5; Length 7736;
Best Local Similarity 54.7%; Pred. No. 2.9;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

OY 110 GTAAGGGGATGGGGGCTAAGAAGTCATTCCTTCTTCCCTGCGGCGACCGG 169
DB 5184 GTTCAGGGATGATGATTAAGACACACAGTGTTCCTCCCGACCGCCAGATG 5125
OY 170 GACCCCTCTCTCCCGGACGATTTCTTCAATATCTTCTTATTCCTAATCCG 229
DB 5124 GAAGTACTCCTCCTCTCCGAGTCTGCTTCCCTCATGCGCTGTGCTCGCTCCC 5065
OY 230 TTGAAGCAACCGCACTA 246
DB 5064 TGGTAGCAGCTGTACCA 5048

RESULT 11
ABS68364/c
ID ABS68364 standard; DNA; 7736 BP.
XX
XX ABS68364;
XX
DT 18-NOV-2002 (first entry)
XX
XX Human DNA-binding protein genomic DNA sequence #67.
XX
XX Human; DNA-binding protein; B cell immunodeficiency; autoimmune disorder;
KW severe combined immunodeficiency; rheumatoid arthritis; Crohn's disease;
KW diabetes mellitus; allergy; asthma; inflammatory condition; thrombosis;
KW graft-versus-host disease; blood-related disorder; atherosclerosis;
KW hyperproliferative disorder; cancer; renal disorder; arrhythmia;
KW acute glomerulonephritis; cardiovascular disorder; respiratory disorder;
KW Goodpasture's syndrome; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; endocrine disorder; Addison's disease;
KW reproductive system disorder; endometriosis; infectious disease;
KW viral infection; bacterial infection; fungal infection; vaccine;
KW

KW gastrointestinal disorder; multiple sclerosis; gene therapy; ds.
XX
XX Homo sapiens.
XX
XX US2002102638-A1.
XX
PD 01-AUG-2002.
XX
XX 17-JAN-2001; 2001US-00764846.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225470P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
XX (RUBEN/) RUBEN S M.
XX (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2002-690611/74.
XX
XX Novel DNA-binding protein useful for diagnosis, prognosis, prevention and
PT treatment of immune, hyperproliferative, respiratory, cardiovascular,
PT reproductive, endocrine, gastrointestinal and neurological disorders.
XX

PS Claim 1; SEQ ID NO 333; 225bp; English.
XX The present invention relates to a new DNA-binding protein. The invention
CC is useful in treating, preventing, diagnosing and/or prognosing
CC immunodeficiencies (e.g. B cell immunodeficiencies, severe combined
CC immunodeficiencies), autoimmune disorders (rheumatoid arthritis, multiple
CC sclerosis, diabetes mellitus), allergic reactions and conditions (e.g.
CC asthma), inflammatory conditions, graft-versus-host disease, blood-
CC related disorders (thrombosis, atherosclerosis), hyperproliferative
CC disorders (e.g. cancer), renal disorders (e.g. acute glomerulonephritis),
CC cardiovascular disorders (e.g. arrhythmia), respiratory disorders
CC (Goodpasture's syndrome), neurological disorders (e.g. Alzheimer's
CC disease, Parkinson's disease), endocrine disorders (e.g. Addison's
CC disease), reproductive system disorders (e.g. endometriosis), infectious
CC diseases (e.g. viral, bacterial or fungal infections) and
CC gastrointestinal disorders (e.g. Crohn's disease). The invention is also
CC useful to stimulate neuronal growth and treat, prevent, and/or diagnose
CC neuronal damage which occurs in certain neuronal disorders or neuro-
CC degenerative conditions. The present nucleic acid sequence represents a
CC human DNA-binding protein genomic DNA sequence of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at <http://seqdata.uspto.gov/sequence>
XX
SQ Sequence 7736 BP; 1672 A; 1981 C; 2164 G; 1919 T; 0 U; 0 Other;
Query Match 5.2%; Score 37.8; DB 6; Length 7736;
Best Local Similarity 54.7%; Pred. No. 2.9;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 110 GTAAGGGGATGAGGAGGAGTCAATTCATCTTTCTTTCCCTGCGGCGGACCCG 169
DB 5184 GTTCAGGGGATGATGTAAGCACACACACAGTGTTCCTCCCAAGCCCAAGATGTC 5125
QY 170 GAAACCCCTCTCTCCCGCACGATTTCTTCCTTCAATCTTCCTTTATCTTATCCCG 229
DB 5124 GAAGTACTCCACTCTCTCCCGAGTGCCTTCCCTCAGGCGCTGACCTCGTCCCG 5065
QY 230 TTGAAGCAACCGACCTA 246
DB 5064 TGGTAGCAGCTGTACCA 5048
RESULT 12
ADC25486/c
ID ADC25486 standard; cDNA; 7736 BP.
XX
AC ADC25486;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human cDNA from extracellular matrix gene 78 #3.
XX
XX Extracellular matrix protein; cytosolic; antibacterial; virucide;
XX neuroprotective; gynaecological; gastrointestinal; cardiac;
XX cardiovascular; Gen; nephrological; antiinflammatory; muscular; Gen;
XX respiratory; Gen; immunosuppressive; cerebroprotective; vasotropic;
XX neotropic; anti-allergic; cancer; bacterial infection; viral infection;
XX neural disorder; immune system disorder; blood disorder;
XX muscular disorder; reproductive disorder; gastrointestinal disorder;
XX pulmonary disorder; cardiovascular disorder; renal disorder;
XX inflammatory disorder; proliferative disorder; human; gene therapy; ss;
XX gene.
XX
XX Homo sapiens.
OS
XX
PN US2003049650-A1.
XX
PD 13-MAR-2003.
XX
XX 07-MAR-2002; 2002US-00091483.
PF
XX 31-JAN-2000; 2000US-0179065P.
PR

PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205155P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0233968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR

PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241851P.
PR 20-OCT-2000; 2000US-0241866P.
PR 20-OCT-2000; 2000US-0241876P.
PR 20-OCT-2000; 2000US-0241889P.
PR 20-OCT-2000; 2000US-0241899P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251888P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764846.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Ruben SM, Barash SC;
XX MPI; 2003-605749/57.
DR P-PSDB; ADC25241.
XX
XX
PT New DNA-binding proteins and gene encoding them, useful for diagnosing,
treating and/or preventing e.g. neurological, inflammatory, infectious,

PT cardiovascular, autoimmune, respiratory, neoplastic or digestive
PT diseases.
XX
XX
PS Disclosure; SEQ ID NO 333; 226pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (cDNA)
CC encoding a human extracellular matrix protein, representing one of 161
CC novel genes. Also included are recombinant vectors, host cells
CC (expressing the protein), the extracellular matrix proteins (including
CC their fragments, epitopes and homologues), an isolated antibody that
CC binds specifically to the protein, diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the
CC presence or absence of a mutation in the nucleic acid and diagnosing a
CC condition based on the presence or absence of the mutation), diagnosing a
CC pathological condition or susceptibility to a pathological condition
CC (comprising determining the presence or amount of expression of the
CC protein in a biological sample and diagnosing a condition based on the
CC presence or amount of expression of the protein), preventing, treating or
CC ameliorating a medical condition by administering the nucleic acid or
CC protein to a mammalian subject, identifying a binding partner to the
CC protein, the gene corresponding to the cDNA sequence, and identifying an
CC activity in a biological assay (comprising expressing the nucleic acid in
CC a cell, isolating the supernatant, detecting an activity in a biological
CC assay and identifying the protein in the supernatant having the
CC activity). The nucleic acids and proteins display the following
CC activities: Cytostatic, antibacterial, Virucide, Neuroprotective.

Query Match 5.2%; Score 37.8; DB 10; Length 7736;
Best Local Similarity 54.7%; Pred. No. 2.9;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 110 GTAAGGGGATGGGGCTAAGAAGTCACTTCTTTCCCTTGGGGTCCGAGCCG 169
Db 5184 GTTCAGGGATGATGATGAACACACACAGTTGTTCCCCACAGCGCCAGATGTG 5125
QY 170 GGACCCCTCCTCTCCCGCAGATTTCTTCTTCATATCTTCTTATTCATCCG 229
Db 5124 GAAGTACTCACTCTCTCCGAGCTGCTTCCCTTCAGCCCTGACCTCGTCC 5065
QY 230 TTGAAGCAACGCACTA 246
Db 5064 TGGTACGAGCTGTACCA 5048

RESULT 13
ABV04394
ID ABV04394 standard; cDNA; 386 BP.
XX
AC ABV04394;
XX
DT 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 4385.
DE
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JR;
XX WPI, 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1, Page 761, 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or incidence of prostate cancer in a patient
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 386 BP, 108 A; 74 C; 74 G; 128 T; 0 U; 2 Other;
Query Match 5.2%; Score 37.4; DB 5; Length 386;
Best Local Similarity 64.4%; Pred. No. 0.95; Mismatches 0; Gaps 0;
Matches 56; Conservative 0; Indels 31;
Qy 628 ACTATGACAGTTCTGGACGCTTAACCTTTATGTCAGAGACTATCAATCATACA 687
Db 175 ATATGTGGAACCTGCTGATTTTTCAGGTTCTCTTTAAGTACTAAGGGAACATCCA 234
Qy 688 GATATTTGTCAAAAAAGACTAA 714
Db 235 GATACCTTCAAAAAAAGCTAA 261
RESULT 14
ID AAL35477/c
XX AAL35477 standard; cDNA; 366 BP.
AC AAL35477;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 819.
XX
KM Cytostatic; immunosuppressive; nocitropic; neuroprotective; antiviral;
KM anti-allergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
KM vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KM neurological disease; infection; human; secreted protein;
KM musculoskeletal system; ss.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001338.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0225799P.
PR 22-AUG-2000; 2000US-0226881P.
PR 22-AUG-2000; 2000US-0226882P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-451937/48.
XX P-PSDB; ABB03895.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the musculoskeletal system including musculoskeletal
PT cancers and also for testing and detection e.g. diagnosis.
XX
XX Claim 1; SEQ ID NO 819; 781bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL3469-AAL3766) and proteins
CC (ABB03087-ABB04109) associated with the musculoskeletal system useful for
CC preventing, treating or ameliorating medical conditions e.g. by protein
CC or gene therapy. The genes are isolated from a range of human tissues
CC disclosed in the specification. The nucleic acids, proteins, antibodies

CC and (ant)agonists are useful in the diagnosis, treatment and prevention
CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 366 BP; 120 A; 59 C; 53 G; 133 T; 0 U; 1 Other;
SQ
Query Match 5.1%; Score 37; DB 4; Length 366;
Best Local Similarity 50.6%; Pred. No. 1.2;
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 549 AGAGATTAGCAGCGGAAAAAACTATGCTAGCTGGAGCTGTTTTCATCATATATA 608
DB 205 AGATTATTTTAAATGAATAATAAATGCGCTTTTAAATGCTTAATTTTAAA 146
QY 609 AAGGAGAAATGTTGCTCCTCATATGTACAGTTTCTGGAGCTTAACTTTTATTCAG 668
DB 145 GAAAGACCGAGTGTCTCTCTATGCTGCAACCAATTCGAATGTGGCTTTTAAAGTAA 86
QY 669 AGGACTATCAATCATCATAGATATTGTCAAAAAAAGACATTAATATATA 722
DB 85 AACTTTTCAATTAANAGATCATTTATACGTTAAAAAAAATCCAGATTGATA 32
RESULT 15
ID AEXS8465/C
ABX58465 standard; CDNA, 366 BP.
AC AEXS8465;
XX
XX DT 26-FEB-2003 (first entry)
XX
XX DE CDNA encoding novel human musculoskeletal system antigen #809.
XX
XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
KW re-vascularisation; thrombosis; arteriosclerosis; mineral content;
KW cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
KW post-operative tissue repair; limb regeneration; neuronal growth;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW AIDS-related complex; chondrocyte growth; bone regeneration;
KW periodontal regeneration; tissue transport; bone graft; skin aging;
KW keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
KW cell growth; organ transplant; cell differentiation; body height; weight;
KW hair colour; eye colour; skin; percentage of adipose tissue;
KW pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
KW depression; tendency for violence; pain; reproductive capability;
KW hormone level; endocrine level; appetite; libido; memory; stress;
KW storage capability; fat content; lipid content; protein content;
KW carbohydrate content; vitamin content; cofactor content;
KW nutritional component.
OS Homo sapiens.
XX
XX PN US2002147140-A1.
XX
XX PD 10-OCT-2002.
XX
XX PF 17-JAN-2001; 2001US-00764877.
XX
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.

CC	PR	11-JUL-2000;	2000US-0217487P.
CC	PR	11-JUL-2000;	2000US-0217496P.
CC	PR	14-JUL-2000;	2000US-0218290P.
CC	PR	26-JUL-2000;	2000US-0220963P.
CC	PR	26-JUL-2000;	2000US-0220964P.
CC	PR	14-AUG-2000;	2000US-0224518P.
CC	PR	14-AUG-2000;	2000US-0224519P.
CC	PR	14-AUG-2000;	2000US-0225267P.
CC	PR	14-AUG-2000;	2000US-0225268P.
CC	PR	14-AUG-2000;	2000US-0225270P.
CC	PR	14-AUG-2000;	2000US-0225477P.
CC	PR	14-AUG-2000;	2000US-0225757P.
CC	PR	14-AUG-2000;	2000US-0225758P.
CC	PR	22-AUG-2000;	2000US-0226688P.
CC	PR	30-AUG-2000;	2000US-0228924P.
CC	PR	01-SEP-2000;	2000US-0229287P.
CC	PR	01-SEP-2000;	2000US-0229343P.
CC	PR	01-SEP-2000;	2000US-0229344P.
CC	PR	01-SEP-2000;	2000US-0229345P.
CC	PR	05-SEP-2000;	2000US-0229509P.
CC	PR	05-SEP-2000;	2000US-0229513P.
CC	PR	08-SEP-2000;	2000US-0231133P.
CC	PR	21-SEP-2000;	2000US-0234223P.
CC	PR	21-SEP-2000;	2000US-0234274P.
CC	PR	25-SEP-2000;	2000US-0234597P.
CC	PR	27-SEP-2000;	2000US-0235834P.
CC	PR	29-SEP-2000;	2000US-0236327P.
CC	PR	29-SEP-2000;	2000US-0236367P.
CC	PR	29-SEP-2000;	2000US-0236368P.
CC	PR	29-SEP-2000;	2000US-0236369P.
CC	PR	29-SEP-2000;	2000US-0236370P.
CC	PR	02-OCT-2000;	2000US-0236802P.
CC	PR	02-OCT-2000;	2000US-0237037P.
CC	PR	02-OCT-2000;	2000US-0237038P.
CC	PR	02-OCT-2000;	2000US-0237039P.
CC	PR	02-OCT-2000;	2000US-0237040P.
CC	PR	13-OCT-2000;	2000US-0239335P.
CC	PR	20-OCT-2000;	2000US-0240960P.
CC	PR	20-OCT-2000;	2000US-0241785P.
CC	PR	20-OCT-2000;	2000US-0241809P.
CC	PR	01-NOV-2000;	2000US-0244617P.
CC	PR	17-NOV-2000;	2000US-0249299P.
CC	PR	08-DEC-2000;	2000US-0251856P.
CC	PR	08-DEC-2000;	2000US-0251868P.
CC	PR	08-DEC-2000;	2000US-0251869P.
CC	XX		
CC	PA	(ROSE/) ROSEN C A.	
CC	PA	(RUBE/) RUBEN S M.	
CC	PA	(BARA/) BARASH S C.	
CC	PI	Rosen CA, Ruben SM, Barash SC,	
CC	PI	WPI, 2003-128199/12.	
CC	DR	P-PSDB; AB013189.	
CC	XX		
CC	XX		
CC	PT	Isolated nucleic acid molecules encoding musculoskeletal system	
CC	PT	associated polypeptides, useful for detecting disorders, e.g. cancer.	
CC	XX		
CC	PS	Claim 1, SEQ ID NO 819; 321pp; English.	
CC	XX		
CC	XX		
CC	CC	The invention describes an isolated nucleic acid molecule comprising a	
CC	CC	sequence encoding musculoskeletal system associated polypeptides useful	
CC	CC	for detecting disorders, e.g., cancer or cancer metastases, in animals or	
CC	CC	humans. The nucleic acid, stimulates re-vascularization of ischaemic	
CC	CC	tissues associated with conditions such as thrombosis, arteriosclerosis,	
CC	CC	and other cardiovascular conditions; treats wounds due to injuries,	
CC	CC	burns, post-operative tissue repair, and ulcers; stimulates angiogenesis	
CC	CC	and limb regeneration; stimulates neuronal growth; can treat and prevent	
CC	CC	neuronal damage occurring in certain disorders or neurodegenerative	
CC	CC	conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-	
CC	CC	related complex; stimulates chondrocyte growth, thus they can be used to	
CC	CC	enhance bone and periodontal regeneration and aid in tissue transports or	
CC	CC	bone grafts/prevents skin aging due to sunburn by stimulating	

CC keratinocyte growth; prevents hair loss, since FGF family members
CC activate hair-forming cells and promotes melanocyte growth; stimulates
CC growth and differentiation of hematopoietic cells and bone marrow cells
CC when used in combination with other cytokines; maintains organs before
CC transplantation or for supporting cell culture of primary tissues;
CC induces tissue of mesodermal origin to differentiate in early embryos;
CC increases or decreases the differentiation or proliferation of embryonic
CC stem cells, besides, haematopoietic lineage; modulates mammalian
CC characteristics, such as, body height, weight, hair colour, eye colour,
CC skin, percentage of adipose tissue, pigmentation, size, and shape (e.g.,
CC cosmetic surgery); modulates mammalian metabolism; changes mammal's metal
CC state or physical state by influencing biophysical, cardiac rhythm,
CC depression, tendency for violence, tolerance for pain, libido, memory,
CC capabilities, hormonal or endocrine levels, appetite, lipid, memory, or
CC stress; increases or decreases storage capabilities, fat content, lipid,
CC protein, carbohydrate, vitamins, minerals, cofactors or other nutritional
CC components. This sequence encodes a novel human musculoskeletal system
CC antigen. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID=20020147140

SQ Sequence 366 BP; 120 A; 59 C; 53 G; 133 T; 0 U; 1 Other;

Query Match 5.1%; Score 37; DB 8; Length 366;

Best Local Similarity 50.6%; Pred. No. 1.2;
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

549 AGAGATTAGCAAGCGGAAAACTATGCTAGCTGGAGTTGTTTTCAATCATATAA 608

Db 205 AGATTATTTTAAATGAAATAAATAAATGGCGCTTTTAATAGTTGCTTAATTTGAAA 146

609 AAGGAGAAATTGTTGCTCACTATGTGACAGTTCTGGGACGCTTAACCTTTATTGCAG 668

Db 145 GAAGGACCAGTGTCTTCTCTATGCTGCACCAATTCCAAATGTTGGCTTTTAAAGTAA 86

669 AGACTATCAATCATACAGATATTGTCAAAAAAAAAAGACTAATAACA 722

Db 85 AACTCTTCAATAANACAGTCATTATACCGTTAAAAAAATCCAGATTGATA 32

Search completed: August 27, 2005, 08:07:49
Job time : 534 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 09:08:05 ; Search time 656 Seconds
(without alignments)
7211.363 Million cell updates/sec

Title: US-10-776-213-2

Sequence: 1 ctttcgattagcagcagcacac.....aaaagacataataacat 723

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCF_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

22: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

23: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

24: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	723	100.0	723	19	US-10-776-213-2
2	723	100.0	11437	19	US-10-776-213-24
3	723	100.0	13073	19	US-10-776-213-20
4	720.4	99.6	850	19	US-10-776-213-30
5	497.4	68.8	680	19	US-10-451-467A-317
6	41	5.7	1024	14	US-10-123-155-198
7	41	5.7	1024	15	US-10-146-731-198

C	8	41	5.7	1024	15	US-10-140-472-198	Sequence 198, App
C	9	41	5.7	1024	15	US-10-141-761-198	Sequence 198, App
C	10	41	5.7	1024	16	US-10-142-885-198	Sequence 198, App
C	11	41	5.7	1024	16	US-10-158-790-198	Sequence 198, App
C	12	41	5.7	1024	17	US-10-137-871-198	Sequence 198, App
C	13	41	5.7	1024	17	US-10-140-923-198	Sequence 198, App
C	14	41	5.7	1024	17	US-10-141-756-198	Sequence 198, App
C	15	41	5.7	1024	17	US-10-141-759-198	Sequence 198, App
C	16	41	5.7	1024	17	US-10-140-805-198	Sequence 198, App
C	17	41	5.7	1024	17	US-10-140-864-198	Sequence 198, App
C	18	41	5.7	1024	18	US-10-142-446-198	Sequence 198, App
C	19	41	5.7	2561	17	US-10-197-824-20	Sequence 20, App1
C	20	39.4	5.4	409	18	US-10-424-599-136845	Sequence 23, App1
C	21	38.6	5.3	2311	19	US-10-466-531-23	Sequence 136845,
C	22	38.4	5.3	544	14	US-10-066-543-1775	Sequence 1775, App
C	23	38.4	5.3	1225	21	US-10-956-157-278	Sequence 278, App
C	24	38.4	5.3	1225	21	US-10-956-157-278	Sequence 278, App
C	25	38.4	5.3	1225	21	US-10-956-157-5512	Sequence 5512, App
C	26	38.4	5.3	1225	21	US-10-956-157-5513	Sequence 5513, App
C	27	38.4	5.3	2355	18	US-10-424-599-29921	Sequence 29921, App
C	28	38.2	5.3	1635	18	US-10-424-599-129944	Sequence 129944,
C	29	37.8	5.2	7736	9	US-09-764-846-333	Sequence 333, App
C	30	37.8	5.2	7736	14	US-10-091-483-333	Sequence 333, App
C	31	37.4	5.2	386	20	US-10-357-930-4385	Sequence 4385, App
C	32	37.4	5.2	436	18	US-10-424-599-96559	Sequence 96559, App
C	33	37.2	5.1	685	13	US-10-027-632-106894	Sequence 106894,
C	34	37.2	5.1	685	13	US-10-027-632-106895	Sequence 106895,
C	35	37.2	5.1	685	17	US-10-027-632-106894	Sequence 106894,
C	36	37.2	5.1	685	17	US-10-027-632-106895	Sequence 106895,
C	37	37.2	5.1	704	13	US-10-027-632-25182	Sequence 25182, App
C	38	37.2	5.1	704	13	US-10-027-632-25183	Sequence 25183, App
C	39	37.2	5.1	704	17	US-10-027-632-25182	Sequence 25182, App
C	40	37.2	5.1	704	17	US-10-027-632-25183	Sequence 25183, App
C	41	37.2	5.1	1121	18	US-10-424-599-94399	Sequence 94399, App
C	42	37	5.1	366	9	US-09-764-877-819	Sequence 819, App
C	43	37	5.1	366	17	US-10-242-515-819	Sequence 819, App
C	44	37	5.1	390	9	US-09-960-352-1976	Sequence 1976, App
C	45	37	5.1	400	11	US-09-969-034-2883	Sequence 2883, App

ALIGNMENTS

RESULT 1

US-10-776-213-2

Sequence 2, Application US/10776213

Publication No. US20040142478A1

GENERAL INFORMATION:

APPLICANT: Astrazeneca AB

TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for Controlling Nucleic Acid Expression

FILE REFERENCE: 3526.82543

CURRENT APPLICATION NUMBER: US/10/776, 213

CURRENT FILING DATE: 2004-02-12

NUMBER OF SEQ ID NOS: 32

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 723

TYPE: DNA

ORGANISM: Saccharomyces cerevisiae

US-10-776-213-2

Query Match	100.0%	Score 723;	DB 19;	Length 723;
Best Local Similarity	100.0%	Pred. No. 1.1e-216;		
Matches 723;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CTTTCGATTAGCAGCAGACATCATGAGCTGCTATATAATACCTACGAGAAA	60	
DB	1	CTTTCGATTAGCAGCAGACATCATGAGCTGCTATATAATACCTACGAGAAA	60	
QY	61	ACCATTAAGAGCAAGCGATTACTTCTGGAAGAAAGAGCAGCGCTTTGAAGGGGAT	120	
DB	61	ACCATTAAGAGCAAGCGATTACTTCTGGAAGAAAGAGCAGCGCTTTGAAGGGGAT	120	

QY 121 GGGGGCTAAGAGTCATTCATCTTTCCCTTCGGGTCGGACCCCGGACCCCTCT 180
DB 121 GGGGGCTAAGAGTCATTCATCTTTCCCTTCGGGTCGGACCCCGGACCCCTCT 180
QY 181 CTCCTCCGACGATTTCTTCTTTCAATCTCTCTTTATTTCTATCCCGTTGAAGAAC 240
DB 181 CTCCTCCGACGATTTCTTCTTTCAATCTCTCTTTATTTCTATCCCGTTGAAGAAC 240
QY 241 GCACTATGACTAATAGTGTGTGACATCTCAATGGCTGTGACTGTGTATCTCACT 300
DB 241 GCACTATGACTAATAGTGTGTGACATCTCAATGGCTGTGACTGTGTATCTCACT 300
QY 301 GGTAAAGGACCGTGGCTCGAAACGGTTCTTCTGTGACATTTCTAGAACAGGGGCTACA 360
DB 301 GGTAAAGGACCGTGGCTCGAAACGGTTCTTCTGTGACATTTCTAGAACAGGGGCTACA 360
QY 361 GTCTCGATTAATAGAAATTAAGGACATTTTGTGTAGCCGCCCGCGCCCTTTCCCA 420
DB 361 GTCTCGATTAATAGAAATTAAGGACATTTTGTGTAGCCGCCCGCGCCCTTTCCCA 420
QY 421 ATAGGAGGCGGAGTTATCGGCGAGCTCTACTTCTCTATTGGGTAAGCCCTTTTC 480
DB 421 ATAGGAGGCGGAGTTATCGGCGAGCTCTACTTCTCTATTGGGTAAGCCCTTTTC 480
QY 481 TGTTTTCGGCCAGTGTGTCTGTGAGGCTGCGCGAGAACATATGATAGGAGATGTAC 540
DB 481 TGTTTTCGGCCAGTGTGTCTGTGAGGCTGCGCGAGAACATATGATAGGAGATGTAC 540
QY 541 TTTTCGATGAGAAATTTAGCAAGCGGAAAAAATATAGCTAGCTGGAGTTGTTTTCAA 600
DB 541 TTTTCGATGAGAAATTTAGCAAGCGGAAAAAATATAGCTAGCTGGAGTTGTTTTCAA 600
QY 601 TCATATTAAGGAGAAATTTGTCTCACTATGTGACAGTTCTGGGAGCTTTACTTT 660
DB 601 TCATATTAAGGAGAAATTTGTCTCACTATGTGACAGTTCTGGGAGCTTTACTTT 660
QY 661 TATTGCAAGAGCTATCAATCATACAGATTTGTCAAAAAAAGACTATATATTA 720
DB 661 TATTGCAAGAGCTATCAATCATACAGATTTGTCAAAAAAAGACTATATATTA 720
QY 721 CAT 723
DB 721 CAT 723

RESULT 2
US-10-776-213-24
; Sequence 24, Application US/10776213
; Publication No. US20040142478A1
; GENERAL INFORMATION:
; APPLICANT: AstraZeneca AB
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for Controlling Nucleic Acid Expression
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/10/776,213
; CURRENT FILING DATE: 2004-02-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 11427
; TYPE: DNA
; ORGANISM: *Saccharomyces cerevisiae*
US-10-776-213-24

Query Match 100.0%; Score 723; DB 19; Length 11427;
Best Local Similarity 100.0%; Pred. No. 4,7e-216;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTGATTAAGACGACACATCACTAGACTGCGTCAATAAAATACACTAGCGAAAA 60
DB 15 CTTTGATTAAGACGACACATCACTAGACTGCGTCAATAAAATACACTAGCGAAAA 74

QY 61 ACCATMAAGAGCAAAAGCAATCTAATTGAAAGAAAGAGCAGCTTTGAAGGGAGAT 120
DB 75 ACCATMAAGAGCAAAAGCAATCTAATTGAAAGAAAGAGCAGCTTTGAAGGGAGAT 134
QY 121 GGGGGCTAAGAGTCATTCATCTTTCCCTTCGGGTCGGACCCCGGACCCCTCT 180
DB 121 GGGGGCTAAGAGTCATTCATCTTTCCCTTCGGGTCGGACCCCGGACCCCTCT 180
QY 181 CTCCTCCGACGATTTCTTCTTTCAATCTCTCTTTATTTCTATCCCGTTGAAGAAC 240
DB 181 CTCCTCCGACGATTTCTTCTTTCAATCTCTCTTTATTTCTATCCCGTTGAAGAAC 240
QY 241 GCACTATGACTAATAGTGTGTGACATCTCAATGGCTGTGACTGTGTATCTCACT 300
DB 241 GCACTATGACTAATAGTGTGTGACATCTCAATGGCTGTGACTGTGTATCTCACT 300
QY 301 GGTAAAGGACCGTGGCTCGAAACGGTTCTTCTGTGACATTTCTAGAACAGGGGCTACA 360
DB 301 GGTAAAGGACCGTGGCTCGAAACGGTTCTTCTGTGACATTTCTAGAACAGGGGCTACA 360
QY 361 GTCTCGATTAATAGAAATTAAGGACATTTTGTGTAGCCGCCCGCGCCCTTTCCCA 420
DB 361 GTCTCGATTAATAGAAATTAAGGACATTTTGTGTAGCCGCCCGCGCCCTTTCCCA 420
QY 421 ATAGGAGGCGGAGTTATCGGCGAGCTCTACTTCTCTATTGGGTAAGCCCTTTTC 480
DB 421 ATAGGAGGCGGAGTTATCGGCGAGCTCTACTTCTCTATTGGGTAAGCCCTTTTC 480
QY 481 TGTTTTCGGCCAGTGTGTCTGTGAGGCTGCGCGAGAACATATGATAGGAGATGTAC 540
DB 481 TGTTTTCGGCCAGTGTGTCTGTGAGGCTGCGCGAGAACATATGATAGGAGATGTAC 540
QY 541 TTTTCGATGAGAAATTTAGCAAGCGGAAAAAATATAGCTAGCTGGAGTTGTTTTCAA 600
DB 541 TTTTCGATGAGAAATTTAGCAAGCGGAAAAAATATAGCTAGCTGGAGTTGTTTTCAA 600
QY 601 TCATATTAAGGAGAAATTTGTGTCTCACTATGTGACAGTTCTGGGAGCTTTACTTT 660
DB 601 TCATATTAAGGAGAAATTTGTGTCTCACTATGTGACAGTTCTGGGAGCTTTACTTT 660
QY 661 TATTGCAAGAGCTATCAATCATACAGATTTGTCAAAAAAAGACTATATATTA 720
DB 661 TATTGCAAGAGCTATCAATCATACAGATTTGTCAAAAAAAGACTATATATTA 720
QY 721 CAT 723
DB 721 CAT 723
QY 735 CAT 737
DB 735 CAT 737

RESULT 3
US-10-776-213-20
; Sequence 20, Application US/10776213
; Publication No. US20040142478A1
; GENERAL INFORMATION:
; APPLICANT: AstraZeneca AB
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for Controlling Nucleic Acid Expression
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/10/776,213
; CURRENT FILING DATE: 2004-02-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 13073
; TYPE: DNA
; ORGANISM: *Saccharomyces cerevisiae*
US-10-776-213-20

Query Match 100.0%; Score 723; DB 19; Length 13073;
Best Local Similarity 100.0%; Pred. No. 5.1e-216;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTGATTAAGACGACACATCACTAGACTGCGTCAATAAAATACACTAGCGAAAA 60

```
|||||
Db 16 CTTTCATTAGACGACACACATCAATAGACTGCGTCAATAAATACTACGAGAAA 75
Qy 61 ACCATAAAGGCAAGAGATACCTACTGTGAAAGAAAAGAGACCCCTTGTAAGGGGAT 120
Db 76 ACCATAAAGGCAAGAGATACCTACTGTGAAAGAAAAGAGACCCCTTGTAAGGGGAT 135
Qy 121 GGGGGCTAAGAAATCACTTCTTTTCCCTTCCGCGGTCCGAGACCCCGGACCCCTCT 180
Db 136 GGGGGCTAAGAAATCACTTCTTTTCCCTTCCGCGGTCCGAGACCCCGGACCCCTCT 195
Qy 181 CTCCTCCGACAGATTTCTTCTTCAATCTTCTTTTATTCCTATCCGTTGAGAACCC 240
Db 196 CTCCTCCGACAGATTTCTTCTTCAATCTTCTTTTATTCCTATCCGTTGAGAACCC 255
Qy 241 GCACATATGACTAATAGTGTGAGACATCTTCAGATGCGTGTGATCTGTATCTCAAGT 300
Db 256 GCACATATGACTAATAGTGTGAGACATCTTCAGATGCGTGTGATCTGTATCTCAAGT 315
Qy 301 GGTAAAGGCAAGAGATACCTACTGTGAAAGAAAAGAGACCCCTTGTAAGGGGAT 360
Db 316 GGTAAAGGCAAGAGATACCTACTGTGAAAGAAAAGAGACCCCTTGTAAGGGGAT 375
Qy 361 GTCCTGATATAGATTAATTAAGGCAATTTTGTGACGCGCGCGCGCGCTTTCCCA 420
Db 376 GTCCTGATATAGATTAATTAAGGCAATTTTGTGACGCGCGCGCGCGCTTTCCCA 435
Qy 421 ATTAGGAGGCGGAGTTTATTCGCGGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 480
Db 436 ATTAGGAGGCGGAGTTTATTCGCGGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 495
Qy 481 TGTTTTCGGCGAGTGTGCTGAGGCTGCGCGGAGAAATAGATTAAGGATTAAC 540
Db 496 TGTTTTCGGCGAGTGTGCTGAGGCTGCGCGGAGAAATAGATTAAGGATTAAC 555
Qy 541 TTTGATGAGAGATTAAGCAAGCGGAAAAAACTATGCTAGCTGAGATTTTTCAA 600
Db 556 TTTGATGAGAGATTAAGCAAGCGGAAAAAACTATGCTAGCTGAGATTTTTCAA 615
Qy 601 TCATATTAAGGAGAAATTTGCTCACTATGTGACATTTTGGAGGCTTAACTTT 660
Db 616 TCATATTAAGGAGAAATTTGCTCACTATGTGACATTTTGGAGGCTTAACTTT 675
Qy 661 TATTGAGAGGACTATCAATCATACAGATTTGTCAAAAAAAGAGCTATATATA 720
Db 676 TATTGAGAGGACTATCAATCATACAGATTTGTCAAAAAAAGAGCTATATATA 735
Qy 721 CAT 723
Db 736 CAT 738

RESULT 4
US-10-776-213-30
; Sequence 30, Application US/10776213
; Publication No. US20040142478A1
; GENERAL INFORMATION:
; APPLICANT: Astrazeneca AB
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for Controlling Nucleic Acid Expression
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/10/776,213
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-776-213-30

Query Match 99.6%; Score 720.4; DB 19; length 850;
Best Local Similarity 99.9%; Pred. No. 7,6e-216;
```

```
Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CTTTCATTAGACGACACACATCAATAGACTGCGTCAATAAATACTACGAGAAA 60
Db 78 CTTTCATTAGACGACACACATCAATAGACTGCGTCAATAAATACTACGAGAAA 137
Qy 61 ACCATAAAGGCAAGAGATACCTACTGTGAAAGAAAAGAGACCCCTTGTAAGGGGAT 120
Db 138 ACCATAAAGGCAAGAGATACCTACTGTGAAAGAAAAGAGACCCCTTGTAAGGGGAT 197
Qy 121 GGGGGCTAAGAAATCACTTCTTTTCCCTTCCGCGGTCCGAGACCCCGGACCCCTCT 180
Db 198 GGGGGCTAAGAAATCACTTCTTTTCCCTTCCGCGGTCCGAGACCCCGGACCCCTCT 257
Qy 181 CTCCTCCGACAGATTTCTTCTTCAATCTTCTTTTATTCCTATCCGTTGAGAACCC 240
Db 258 CTCCTCCGACAGATTTCTTCTTCAATCTTCTTTTATTCCTATCCGTTGAGAACCC 317
Qy 241 GCACATATGACTAATAGTGTGAGACATCTTCAGATGCGTGTGATCTGTATCTCAAGT 300
Db 318 GCACATATGACTAATAGTGTGAGACATCTTCAGATGCGTGTGATCTGTATCTCAAGT 377
Qy 301 GGTAAAGGCAAGAGATACCTACTGTGAAAGAAAAGAGACCCCTTGTAAGGGGAT 360
Db 378 GGTAAAGGCAAGAGATACCTACTGTGAAAGAAAAGAGACCCCTTGTAAGGGGAT 437
Qy 361 GTCCTGATATAGATTAATTAAGGCAATTTTGTGACGCGCGCGCGCGCTTTCCCA 420
Db 438 GTCCTGATATAGATTAATTAAGGCAATTTTGTGACGCGCGCGCGCGCTTTCCCA 497
Qy 421 ATTAGGAGGCGGAGTTTATTCGCGGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 480
Db 498 ATTAGGAGGCGGAGTTTATTCGCGGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 557
Qy 481 TGTTTTCGGCGAGTGTGCTGAGGCTGCGCGGAGAAATAGATTAAGGATTAAC 540
Db 558 TGTTTTCGGCGAGTGTGCTGAGGCTGCGCGGAGAAATAGATTAAGGATTAAC 617
Qy 541 TTTGATGAGAGATTAAGCAAGCGGAAAAAACTATGCTAGCTGAGATTTTTCAA 600
Db 618 TTTGATGAGAGATTAAGCAAGCGGAAAAAACTATGCTAGCTGAGATTTTTCAA 677
Qy 601 TCATATTAAGGAGAAATTTGCTCACTATGTGACATTTTGGAGGCTTAACTTT 660
Db 678 TCATATTAAGGAGAAATTTGCTCACTATGTGACATTTTGGAGGCTTAACTTT 737
Qy 661 TATTGAGAGGACTATCAATCATACAGATTTGTCAAAAAAAGAGCTATATATA 720
Db 738 TATTGAGAGGACTATCAATCATACAGATTTGTCAAAAAAAGAGCTATATATA 797
Qy 721 CA 722
Db 798 AA 799

RESULT 5
US-10-451-467A-317
; Sequence 317, Application US/10451467A
; Publication No. US20040161840A1
; GENERAL INFORMATION:
; APPLICANT: CONTERBAS, ROLAND HENRI
; APPLICANT: EBERHARDT, INES
; APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
; APPLICANT: REEKMAN, RIEKA JOSEPHINA
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN YEAST AND FUNGI
; FILE REFERENCE: JAB-1667
; CURRENT APPLICATION NUMBER: US/10/451,467A
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 317
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-451-467A-317

Query Match 99.6%; Score 720.4; DB 19; length 850;
Best Local Similarity 99.9%; Pred. No. 7,6e-216;
```

```

; PRIOR APPLICATION NUMBER: EP 01870003.9
; PRIORITY FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 732
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 317
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-451-467A-317

Query Match
Best Local Similarity 68.8%; Score 497.4; DB 19; Length 680;
Matches 498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 224 ATCCGCTTGAAGAACCGCACTATGATTAATGCTGAGACATCTCATGCTGACT 283
DB 1 ATCCGCTTGAAGAACCGCACTATGATTAATGCTGAGACATCTCATGCTGACT 60
QY 284 TGTGTATCTCACTGCTGTAAGGCAAGGCTGCTGGAAGCGCTTCTTGTGACATTT 343
DB 61 TGTGTATCTCACTGCTGTAAGGCAAGGCTGCTGGAAGCGCTTCTTGTGACATTT 120
QY 344 CTGAACAGGGGCTACAGTCTCGATATATGAATATATAGGCAATTTTGTGCGCGCGC 403
DB 121 CTGAACAGGGGCTACAGTCTCGATATATGAATATATAGGCAATTTTGTGCGCGCGC 180
QY 404 GCGGCGCCGCTTCCCAATAGAGGCGCAGTTTATCGCGAGCTCTACTTCTTCTAT 463
DB 181 GCGGCGCCGCTTCCCAATAGAGGCGCAGTTTATCGCGAGCTCTACTTCTTCTAT 240
QY 464 TTGGGTAAGCCCTTCTGTTTTCGCGCAGTGTCTCTGAGGCTGCGCGGAGAACATA 523
DB 241 TTGGGTAAGCCCTTCTGTTTTCGCGCAGTGTCTCTGAGGCTGCGCGGAGAACATA 300
QY 524 GTGATATAGGAGTATCTTTCATGAGATTTAGCAAGGAGGAGGAGGAGGAGGAGGAG 583
DB 301 GTGATATAGGAGTATCTTTCATGAGATTTAGCAAGGAGGAGGAGGAGGAGGAGGAG 360
QY 584 TGGGAGCTTGTTCATCATATATAAGGAGAGAAATTTGCTCACTATGACAGTTTC 643
DB 361 TGGGAGCTTGTTCATCATATATAAGGAGAGAAATTTGCTCACTATGACAGTTTC 420
QY 644 TGGGAGCTTGTTCATCTTTATGCAAGGAGCTATCAATCATACATATTTGCAAAAAA 703
DB 421 TGGGAGCTTGTTCATCTTTATGCAAGGAGCTATCAATCATACATATTTGCAAAAAA 480
QY 704 AAAAAAGCTATATATATACA 722
DB 481 AAAAAAGCTATATATATATA 499

RESULT 6
US-10-123-155-198/c
; Sequence 198, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
```

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; PRIORITY FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-198

Query Match
Best Local Similarity 5.7%; Score 41; DB 14; Length 1024;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

QY 47 TACACTAGGAAAAACCAATAAGCAAGGATACCTACTTTGGAGAGAAAAAGAGCAG 106
DB 954 T.CRYAN..MMMHV..MWT.HYCD.NYTB..BH.TH.H.BB.SN.S.N..YN.NYHS 895
QY 107 CTGTAAAGGGGATGGGCTAAGAAGCATTCATCTTCTTCCCTTCCGCGTCCGAGC 166
DB 894 GA.T...MM.SBTASD.MKW..WBH.H..SBCAGB.HYBN..YNYSS.WHS..A 835
QY 167 CCGGAGCCCTCCCTCTCCCGCAGATTTCTTCTCATATCTTCTTATTTCTATC 226
DB 834 ST..DGMNC.SGT.K.B.YV...D.RST.B.SCN.NNM.WC.M.RATB.M..R.N..N 775
QY 227 CCGTTGAAGCAACCGCACTATGACTTAATGAGTGTGAGCATCTCATGAGCTGTACTGT 286
DB 774 RN.KTH..MH..R.TYSTTDW...HM..S.RV.....N..RCTYT.S.TTH..CTYNS 715
QY 287 GTGATCTCACAGTGTGTAACGCGACCGGTGCGGAAACGGTTCCTTGACATTTCTA 346
DB 714 BS.B.DNY..H..YMYV.HRNBX.RCA.N.NC..WSCM.RA.YTD.SWMSB..T.S 655
QY 347 GAAACAGGGCTACAGTCTCGATATATAGATATAGCGCATTTTGTAGCGCGCGCG 406
DB 654 SBDYNCB.B.A...W.RSN..M.TWAMTS.HR..D..A..YN.TPANC..A.B.RCK 595
QY 407 GCGCCCGTTCCCAATAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 466
DB 594 ..AM.HXK...NBRNCNM.T.S..ANW.HNYTTCRMD.A.RH..CY..BBDCK.NT.KY 535
QY 467 GTTAAGCCCTTCTGTTTTCGCGCAGTGTGCTGCTGAGGCTGCGCGGAGAACATAGTG 526
DB 534 ..TTCMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KWMC.R.RH 475
QY 527 ATTAAGGATGTACTTTCATGAGAGATTTAGCAAGCGGAGGAGGAGGAGGAGGAGGAG 586
DB 474 ACB...RDT.C.M.WBH.WMB.AB.HC.W.DG..DB.BKABH..RS.SBSB.H..Y..B 415
QY 587 GAGTGTTCATCATATATAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 646
DB 414 ..S.CWN.KTRR.DM.BT..H.M..SR.BM.RH.R...YVA.D.TRRCSCYV..TH.YTR 355
QY 647 GAGCTTAACTTTATGCGAGAGCATATCAATCATACATGATATTTGCAAAAAA 706
DB 354 .YBBSMAAA.YAA.NG.TWT.RRASK..A.S.DHABWCH..BSA.B.KNS.S.RAT. 295
QY 707 AAGACTATATATATAC 721
DB 294 DM.S.M.ATMR...H 280

RESULT 7
US-10-146-731-198/c
; Sequence 198, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
```

```
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C323
CURRENT APPLICATION NUMBER: US/10/146,731
CURRENT FILING DATE: 2002-05-15
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 198
LENGTH: 1024
TYPE: PRT
ORGANISM: Homo Sapien
US-10-146-731-198
```

Query Match 5.7%; Score 41; DB 15; Length 1024;

Best Local Similarity 7.7%; Pred. No. 0.1;

Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

```
QY 47 TACACTACGGAAGAAACGATTAAGAGCAAGCGATCTTCTTGAGAGAAAGAGACAG 106
DB 954 T.CRYAN..MMHY..MMT.HYCD..NYTBD..BH.TH.H.BB..SN.S.N..YN..NYHHS 895
QY 107 CTGTGAAGGGAGTGGGGCTTAAGAGTCATTCATCTTTCTTCCGCGCGCGAGC 166
DB 894 GA.T...MM.SBTASD.MKW..WBH.H..SBCAGR.B.HYBN...YVNS.WHS..A 835
QY 167 CCGGAGCCCTCTCTCCCGCAGCATTTCTTCCATATCTTCTTTATCTTATC 226
DB 834 ST..DGDNMC.SGT.K.B.YY..D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
QY 227 CGGTGAAGCAACCGCATATGCTTAAGTGCTGAGCATCTCCATGCTGAGCTGT 286
DB 774 RN.KYH..MH..R.TYSTIDW..HM..S.RY.....N...RCTYT.S.THH..CTVNS 715
QY 287 GTGTATCTCAGAGTGTAAAGCGACCGGTGCGAAGCGTTCTTCTGACATTTCTA 346
DB 714 BS.B.DNY..H..YVNY.HRNB.Y.RCA.N.NC..WSCMH.RA.YDD..SMNSBW..T.S 655
QY 347 GAACGAGGCTACAGTCTCGATATAGATTAATTAAGCGCATTTTCTAGCGCGCG 406
DB 654 SBDSTNCB.A...W.RSNN..M.TMAMTS..HR..D..A..YN.TAANC..A.B.RCK 595
QY 407 GCGCCGTTTCCCAATAGGAGCGAGTTATCGCGGAGCTCTACTTCTTCTATTG 466
DB 594 ..AM.HKB...NBRMCN.YM.T.S.ANW..HNYTTCMD.A.RH..CY..BSDDK..NT.RY 535
QY 467 GGTAAAGCCCTTCTGTTTTCGCGCAGTGTGCTGAGGCTGCGCGGAGACATATG 526
DB 534 ..TTCMRT.WH.BYH.HAT..B..SB.DSST..CH..KB.D.BHA..Y.M.KW.C.R.RH 475
QY 527 ATTAAGGATGTAACTTTCATGAGAGATTAAGCAAGCGGAAAAAACTATGCTAGCTG 586
DB 474 ACB...RDT.C.M.WBH.WMB..AB.HC.W.DG..DB.BKABH..RS.SBS.H...Y..B 415
QY 587 GAGTGTGTTTTCATATATAAAGGAGAAATTTGCTCACTATGTGACATTTCTG 646
DB 414 ..S.CMN.KTTR..DM..BT..H.M..SR.BM.RH.R...YMA..D..TRHSCY..TH.YTR 355
QY 647 GAGCTTAATCTTTATGAGAGAGCATCAATCATACAGATATGTCAAAAAAA 706
```

```
DB 354 .YBBCSMAAA.YRA.NG.TWT.RRASK..A.S.DHAEBCH...BSA.B.KUS.S.RAT. 295
QY 707 AAGACTAATATATAC 721
DB 294 DM.S.M.ATMR...H 280
```

RESULT 8

US-10-140-472-198/C

Sequence 198, Application US/10140472

Publication No. US2003013888A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerlitsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3330R1C168

CURRENT APPLICATION NUMBER: US/10/140,472

CURRENT FILING DATE: 2002-05-06

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 198

LENGTH: 1024

TYPE: PRT

ORGANISM: Homo Sapien

US-10-140-472-198

Query Match 5.7%; Score 41; DB 15; Length 1024;

Best Local Similarity 7.7%; Pred. No. 0.1;

Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

```
QY 47 TACACTACGGAAGAAACGATTAAGAGCAAGCGATCTTCTTGAGAGAAAGAGACAG 106
DB 954 T.CRYAN..MMHY..MMT.HYCD..NYTBD..BH.TH.H.BB..SN.S.N..YN..NYHHS 895
QY 107 CTGTGAAGGGAGTGGGGCTTAAGAGTCATTCATCTTTCTTCCGCGCGCGAGC 166
DB 894 GA.T...MM.SBTASD.MKW..WBH.H..SBCAGR.B.HYBN...YVNS.WHS..A 835
QY 167 CCGGAGCCCTCTCTCCCGCAGCATTTCTTCCATATCTTCTTTATCTTATC 226
DB 834 ST..DGDNMC.SGT.K.B.YY..D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
QY 227 CGGTGAAGCAACCGCATATGCTTAAGTGCTGAGCATCTCCATGCTGAGCTGT 286
DB 774 RN.KYH..MH..R.TYSTIDW..HM..S.RY.....N...RCTYT.S.THH..CTVNS 715
QY 287 GTGTATCTCAGAGTGTAAAGCGACCGGTGCGAAGCGTTCTTCTGACATTTCTA 346
DB 714 BS.B.DNY..H..YVNY.HRNB.Y.RCA.N.NC..WSCMH.RA.YDD..SMNSBW..T.S 655
QY 347 GAACGAGGCTACAGTCTCGATATAGATTAATTAAGCGCATTTTCTAGCGCGCG 406
DB 654 SBDSTNCB.A...W.RSNN..M.TMAMTS..HR..D..A..YN.TAANC..A.B.RCK 595
QY 407 GCGCCGTTTCCCAATAGGAGCGAGTTATCGCGGAGCTCTACTTCTTCTATTG 466
```

```

Db      594 ..AW.HKB...NBRMCNM.T.S.ANW.HNYTTCNRM.D.A.RH..CY..BSDCK.NT.KY.535
QY      467 GGTAAAGCCCTTTCTGTTTCGCGCAGTGTGCTGCAGAGTGGCGCGGAGAACATAGT 526
Db      534 ..TTCMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KNWC.R.RH.475
QY      527 ATAAGGATGTAACTTTGCATGAGAGATTGCAAGCGGAAAAAACTATGGCTAGCTGG 586
Db      474 ACB...RDT.C.M.WBH.WMB.AB.HC.W.DG..DB.BKABH..RS.SBS.H...Y..B.415
QY      587 GAGTGTCTTTCATCATATATAAAGGAGAAATGTGCTCACTATGTACAGTTTCTGG 646
Db      414 ..S.CWN.KTRR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHSCYY.TH.YTR.355
QY      647 GAGCTTACTTTATTCATGACAGACTATCAATCATCATATTTGCAAAAAAAA 706
Db      354 .YBBBCSMAAA.YRA.NG.TWT.RRASK..A.S.DHABWCH..BSA.B.KNS.S.RAT.295
QY      707 AAGACTAATAATAAC 721
Db      294 DM.S.M.ATMR...H.280

```

RESULT 9

```

US-10-141-761-198/c
; Sequence 198, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-761-198

```

Query Match 5.7%; Score 41; DB 15; Length 1024;

Best Local Similarity 7.7%; Pred. No. 0.1;

Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

```

QY      47 TACACTAAGGAAAAACCATTAAGAGCAAGCGCTACTCTTGGAGGAAAAAGACACAG 106
Db      954 T.CRYAN..MMWHY..MWT.HYCD.NYTB.D..BH.TH.F.BB.SN.S.N..YN.NYHHS.895
QY      107 CTGTGAAGGGGATGGGGGCTAGAAGTCATCTTTCTTCCCTGCGGCTCGGAC.166
Db      894 GA.T...MM.SPTASD.MKW...WBH.H..SBGAGR.BYBN...YVNS.WHS..A.835
QY      167 CCGGACCCCTCTCTCCCGCAGCATTTCTTCTTCAATATCTTCTTATCTTCTTATC 226
Db      834 ST..DGDWNC.SGT.K.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N.775

```

```

QY      227 CCGTTGAAGCAACCGCACTATGACTTAATGTGTGTCGACATCTCCAGCTGTGACTTGT 286
Db      774 RN.KYH..MH..R.TYSTTDW...HM..S.RY.....N...RCTYT.S.THH..CTVNS.715
QY      287 GTGTATCTCAGAGTGTAAAGCGACCCGTGCTCGGAAAACGTTCTTGTGTGACATTTCTA 346
Db      714 BS.B.DNY...H...YVNY..HRNB.Y.RCA.N.NC..WSCMH.RA.YTD.SNWSBW..T.S.655
QY      347 GAACAGGGGCTACAGTCTCGATTAATAGATATAAGCCATTTTGTCTAGCGCCGCCGCG 406
Db      654 SBDSTNCB.A...W.RSNM..M.TMAMTS.HR..D..A..YN.TAANC..A.B.RCK.595
QY      407 GCGCCGTTTCCCATAGAGGAGCGCAGTTTATCGCGGAGCTTACTTCTTCTAATTTG.466
Db      594 ..AW.HKB...NBRMCNM.T.S.ANW.HNYTTCNRM.D.A.RH..CY..BSDCK.NT.KY.535
QY      467 GGTAAAGCCCTTTCTGTTTCGCGCAGTGTGCTGCAGAGGCTGCGCCCGGAGAACATAGT 526
Db      534 ..TTCMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KNWC.R.RH.475
QY      527 ATAAGGATGTAACTTTGCATGAGAGATTGCAAGCGGAAAAAACTATGGCTAGCTGG.586
Db      474 ACB...RDT.C.M.WBH.WMB.AB.HC.W.DG..DB.BKABH..RS.SBS.H...Y..B.415
QY      587 GAGTGTCTTTCATCATATATAAAGGAGAAATGTGCTCACTATGTGACAGTTTCTGG.646
Db      414 ..S.CWN.KTRR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHSCYY.TH.YTR.355
QY      647 GAGCTTACTTTATTCATGACAGACTATCAATCATCATATTTGCAAAAAAAA 706
Db      354 .YBBBCSMAAA.YRA.NG.TWT.RRASK..A.S.DHABWCH..BSA.B.KNS.S.RAT.295
QY      707 AAGACTAATAATAAC 721
Db      294 DM.S.M.ATMR...H.280

```

RESULT 10

```

US-10-142-885-198/c
; Sequence 198, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C248
; CURRENT APPLICATION NUMBER: US/10/142,885
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-885-198

```

```

Query Match      5.7%; Score 41; DB 16; Length 1024;
Best Local Similarity 7.7%; Pred. No. 0.1;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

QY 47 TACACTACGAAAAACCATTAAGAGCAAGGACCTACTTCTGGAAGAAAAAGAGACAG 106
DB 954 T.CRYAN..MMMHY..MWT.HYCD..NYTBD..BH.TH.H.BB..SN.S.N..YN..NYNHS 895
QY 107 CTTGTAAGGGGAGGAGGCTAAGATCATCTTCTTCTTCCCTGCGCGGAC 166
DB 894 GA.T...MM.SBTAASD..MKW..WBH.H..SBCAGR.B..HYBN...YVNS.S..A 835
QY 167 CCGGAGACCCCTCTCTCCCGGACAGATTTCTTCTTCAATCTCTTTATCTTATC 226
DB 834 ST..DGDMMC..SGT.K.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
QY 227 CCGTTGAAGCAACCGCATATGCTAATGCTGCTGACATCTTCATGCTGACCTGT 286
DB 774 RN.KYH..MH..R.TYSTTDM...HM..S.RY.....N...RCTYT.S.THH..CTVNS 715
QY 287 GTGTATCTCAGATGTGAACGACGCGTGGCTGGAAGCGTTCTTCTGACATTTCTA 346
DB 714 BS.B.DNY...H...YWNV..HRNBY..RCA.N.NC..WSCMH..RA.YD..SNMSBW..T.S 655
QY 347 GAACAGGGGCTACAGTCTCGATTAATGAATTAATAGCGCATTTTCTGAGCGCGCG 406
DB 654 SBDSYNCB.A...W.RSNN..M.TMAMTS..HR..D..A..YN.TAANC..A.B.RCK 595
QY 407 GCGCCCGTTTCCCAATAGGAGGCGAGCTTATGCGGCGGCTCTACTTCTTCTTAT 466
DB 594 ..AM.HKB...NBRMCNMY.T.S.ANW..HNYTNCMD.A.RH..CY..BSDCK.NT.KY 535
QY 467 GGTAAAGCCCTTCTGTTTTCGCGCAGTGTGCTGAGGCTGCGCGGAGCAATAGTG 526
DB 534 ..TTCMRT..WH.BYH.HAT..B..SB.DSST..CH..KB.D.BHA..Y.M.KWNC.R.RH 475
QY 527 ATAAAGGATGTAATCTTTCATGAGAGATTAAGAGCGGAAAAAACTATGCTAGCTG 586
DB 474 ACB...RDT.C.M.WBH..WBH.AB.HC.W.DG..DB.BKABH..RS.SBS.H...Y..B 415
QY 587 GAGTTGTTTTCATCATATATAAAGAGAAATGTTGCTCACTATGTGACAGTTTCTG 646
DB 414 ..S.CMN.KTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHSCYV.TH.YTR 355
QY 647 GAGCTTTAATCTTTATGAGAGAGCTATCAATCATACAGATATGTCAAAAAAAA 706
DB 354 ..YBBBSMAA..YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT. 295
QY 707 AAGACTAATATAAC 721
DB 294 DM.S.M.ATMR...H 280

RESULT 11
US-10-158-790-198/c
; Sequence 198, Application US/10158790
; Publication No. US20030180879A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltisen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tuma, Daniel
; APPLICANT: Watanabe, Colin K

```

```

; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C448
; CURRENT APPLICATION NUMBER: US/10/158,790
; PRIOR FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-158-790-198

Query Match      5.7%; Score 41; DB 16; Length 1024;
Best Local Similarity 7.7%; Pred. No. 0.1;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

QY 47 TACACTACGAAAAACCATTAAGAGCAAGGACCTACTTCTGGAAGAAAAAGAGACAG 106
DB 954 T.CRYAN..MMMHY..MWT.HYCD..NYTBD..BH.TH.H.BB..SN.S.N..YN..NYNHS 895
QY 107 CTTGTAAGGGGAGGAGGCTAAGATCATCTTCTTCTTCCCTGCGCGGAC 166
DB 894 GA.T...MM.SBTAASD..MKW..WBH.H..SBCAGR.B..HYBN...YVNS.S..A 835
QY 167 CCGGAGACCCCTCTCTCCCGGACAGATTTCTTCTTCAATCTCTTTATCTTATC 226
DB 834 ST..DGDMMC..SGT.K.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
QY 227 CCGTTGAAGCAACCGCATATGCTAATGCTGCTGACATCTTCATGCTGACCTGT 286
DB 774 RN.KYH..MH..R.TYSTTDM...HM..S.RY.....N...RCTYT.S.THH..CTVNS 715
QY 287 GTGTATCTCAGATGTGAACGACGCGTGGCTGGAAGCGTTCTTCTGACATTTCTA 346
DB 714 BS.B.DNY...H...YWNV..HRNBY..RCA.N.NC..WSCMH..RA.YD..SNMSBW..T.S 655
QY 347 GAACAGGGGCTACAGTCTCGATTAATGAATTAATAGCGCATTTTCTGAGCGCGCG 406
DB 654 SBDSYNCB.A...W.RSNN..M.TMAMTS..HR..D..A..YN.TAANC..A.B.RCK 595
QY 407 GCGCCCGTTTCCCAATAGGAGGCGAGCTTATGCGGCGGCTCTACTTCTTCTTAT 466
DB 594 ..AM.HKB...NBRMCNMY.T.S.ANW..HNYTNCMD.A.RH..CY..BSDCK.NT.KY 535
QY 467 GGTAAAGCCCTTCTGTTTTCGCGCAGTGTGCTGAGGCTGCGCGGAGCAATAGTG 526
DB 534 ..TTCMRT..WH.BYH.HAT..B..SB.DSST..CH..KB.D.BHA..Y.M.KWNC.R.RH 475
QY 527 ATAAAGGATGTAATCTTTCATGAGAGATTAAGAGCGGAAAAAACTATGCTAGCTG 586
DB 474 ACB...RDT.C.M.WBH..WBH.AB.HC.W.DG..DB.BKABH..RS.SBS.H...Y..B 415
QY 587 GAGTTGTTTTCATCATATATAAAGAGAAATGTTGCTCACTATGTGACAGTTTCTG 646
DB 414 ..S.CMN.KTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHSCYV.TH.YTR 355
QY 647 GAGCTTTAATCTTTATGAGAGAGCTATCAATCATACAGATATGTCAAAAAAAA 706
DB 354 ..YBBBSMAA..YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT. 295
QY 707 AAGACTAATATAAC 721
DB 294 DM.S.M.ATMR...H 280

RESULT 12
US-10-137-871-198/c
; Sequence 198, Application US/10137871
; Publication No. US20030207350A1
; GENERAL INFORMATION:

```

```

/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zhenli
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3330R1C153
/ CURRENT APPLICATION NUMBER: US/10/137,871
/ CURRENT FILING DATE: 2002-05-03
/ Prior Application removed - See Palm or File Wrapper
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 198
/ LENGTH: 1024
/ TYPE: PR1
/ ORGANISM: Homo Sapien
/ US-10-137-871-198

```

	Query Match	5.7%	Score 41	DB 17	Length 1024
	Best Local Similarity	7.7%	Pred. No. 0.1		
	Matches 52	Conservative 194	Mismatches 429	Indels 0	Gaps 0
QY	47 TACACTACGAAAAACCATAAAGCAAGGATCACTTCTGGAAGAAAAAGACACG	106			
Db	954 T.CRYAN..MMMHY..MWT.HYCD.NYTB.D..BH.TH.H.BH.SN.S.N..YN..NYHNS	895			
QY	107 CTTGTGAAGGGGAGATGGGGCTTAAGAAGCATTCACCTTTCTTTCCCTTCGGGTCGGAC	166			
Db	894 GA.T...MM.SBPAASD.MKM...WBH.H..SBQAGRB.HYBN...YNNYS.WHS..A	835			
QY	167 CCGGAGCCCTCTCTCTCCCGACAGATTTCTTCCTTATATCTCTTTATTTCCATTC	226			
Db	834 ST..DSDMNC.SGT.K.B.YY..D.RST.B.SCN.NNM.MC.M.RATB.M..R.N..N	775			
QY	227 CCGTTGAAGCAACGCGACTATGACTAAATGGTGCTGACATCTCCAGGCTGTGACTTGT	286			
Db	774 RN.KYH..MH..R.TYSTTIDW...HM.S.RY.....N..RCTYT.S.THH..CTYNS	715			
QY	287 GTGATATCAAGTGGTAAAGGCAACCGGTGGGAAACGGTTCCCTTGATACATTTCTA	346			
Db	714 BS.B.DMYT..H...YMYT.HRNB.I.RCA.N.NC..WSCMI.RA.YTD.SMNSMB..T.S	655			
QY	347 GAACAGGGGGCTACAGTCTCGATTAATAGATTAAGCCGATTTTGTCTAGCGCCGCG	406			
Db	654 SBDSYNCEB.A...W.RSNN..M.TMAMTS.HR..D..A..YN.TAANC..A.B.RCK	595			
QY	407 GCGCCCGTTTCCCAATAGAGAGCGCGAGTTTATCGGCGAGCTACTTTCTTATTTTG	466			
Db	594 .AM.HKB..NBRMCNMY.T.S.AMW.HNYYTCRMD.A.RH..CY..BDDCK.NT.KY	535			
QY	467 GGTAAGCCCTTTCGTTGTTGGCGCAGTGGTGCTGCGAGGCTGGCGGAGAACATAGTG	526			
Db	534 .TTTCMMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KNWC.R.RH	475			
QY	527 ATAAGGATGTAACTTTCGATGAGAGAAATTAAGCAAGGAAAAAACTATGCTAGCTGG	586			
Db	474 ACB...RDT.C.M.WBH.WMB.AB.HC.W.DG..DB.BKABH..RS.SBS.H...Y..B	415			
QY	587 GAGTTGTTTTCAATCATATAAAAGGAGAAATTTGCTGACATGTGACAGTTTCTGG	646			
Db	414 ..S.CMN.KTTR..DM.BT..H.M..SR.BM.RH.R...YNA.D..TRHCSTCY..TH.YTR	355			

[illegible]

```

RESULT 13
; US-10-140-923-198/c
; Sequence 198, Application US/10140923
; Publication No. US20030207355A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geriltson, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C188
; CURRENT APPLICATION NUMBER: US/10/140,923
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-140-923-198

```

	Query March	5.7%; Score 41; DB 17; Length 1024;
	Best Local Similarity	7.7%; Pred. NO. 0.1;
	Matches	52; Conservative 194; Mismatches 429; Indels 0; Gaps 0
QY	47 TACACTACGGAAAAACCATAAAGCAAGAAGCATCTACTTGAAGGAAAAAGACACG	106
	::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	954 T.CRYAN..MMMH..MWT.HYCD.NYTBD..BH.TH.H.BB.SN.S.N..YN..NYHS	895
QY	107 CTGTGAAGGGGGATGGGGCTTAAGAATGCATTCTTTCCCTCGGGGCCGAC	166
	::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	894 GA.T...NM..SBTPASD.MKM...WBH.H..SECAGRB..HYBN...YNNYS.WHS..A	835
QY	167 CCGGAGCCCTCCTCCTCCCCGACAGATTCCTCTTTCATATATCTTCTTTATTCCTATC	226
	::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	834 ST..DEGMNC..SGT.K.B.YY..D.RST..B.SCN.NNM.WC.M.PAATB..M..R.N..N	775
QY	227 CCGTTGAAGCAACCGCACTATGACTAAATGSGTGCGACATCTGCATGGCTGTGACTTGT	286
	::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	774 RN.KYH..MH..R.TVSTTDM...HM..S.RY.....N...RCTYT.S.THH..CTYNS	715
QY	287 GTGATCTTCACAGTGGTGAACGGCACCGGTGGCTCGGAACGGTTCTTCCTGTCACATTTCTA	346
	::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	714 BS..B.DMY...H...YNNY..HRMBY..RCA.N.NC..WSCMH.PA.YDD.SMUSMB..T.S	655
QY	347 GAACAGGGGGCTACAGTCTCGATATAGAAATAAAGGACATTTTGGCTAGCGCCGCGC	406
	::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	654 SBDSYCNBB.A...W..RSNN..M.TPAAMS.HR..D..A...YN.TPANC..A.B.RCK	595


```

QY 407 GCGCCGTTTCCCAATAGGAGCGCATTTATCGCGGAGCTCTACTTTCTTATTTG 466
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 594 ..AM.HKB..NBRMCNYM.T.S.ANW..HNYTTCMD.A.RH..CY..BSDCK.NT.KY 535
QY 467 GGTAAAGCCCTTCTCTGTTTTCGCGCAGTGTCTCTGCGAGCTCGCGCGAACAATAGT 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 534 ..TTCCMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KNWC.R.RH 475
QY 527 ATAGAGGATGTACTTTTCATGTAGAGATATAGCAAGCGGAAAAAACTATGCTACTG 586
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 474 ACB...RDT.C.M.WBH.WWB.AB.HC.W.DG..DB.BKABH..RS.SBS.H...Y..B 415
QY 587 GAGTTGTTTTCATCATATAAAGGAGAAATGTTGCTCACTATGTACAGTTTCTG 646
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 414 ..S.CMN.KTTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHSCYV.TH.YTR 355
QY 647 GAGCTCTTAATCTTTATTTGAGAGAGACTCAATCATACATAGATATTTGCAAAAAA 706
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 ..YBBBCSMABA.YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT. 295
QY 707 AAGACTAATATATAC 721
Db 294 DM.S.M.ATMR...H 280

```

RESULT 14

```

US-10-141-756-198/c
; Sequence 198, Application US/10141756
; Publication No. US20030207359A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C200
; CURRENT APPLICATION NUMBER: US/10/141,756
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-756-198

```

```

Query Match 5.7%; Score 41; DB 17; Length 1024;
Best Local Similarity 7.7%; Pred. No. 0.1;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

```

```

QY 47 TACACTACGAAAAACCATTAAGAGCAAAAGCTACTTCTTGAAGAAAGAGACAG 106
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 954 T.CRYAN..MMHY..MWT.HYCD.NYTB.D..BH.TH.H.BB.SN.S.N..YN..NYHS 895
QY 107 CTGTGAAGGGAGATGGGGCTAAGAAATCATTCATTTCTTTCCCTCGCGGTCCGAG 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 894 GA.T...NM.SPTASD.MKW..WBH.H..SBCAGR.B.HYBN...YNNNS.WHS..A 835
QY 167 CCGGAGCCCTCTCTCCCGGACGATTTCTTCCATATATCTCTTATTTCTATC 226

```

```

Db 834 ST..DGDMMC.SGT.K.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
QY 227 CCGTTGAAGCAACCGACATATGACTTAATGCTGCTGAGACATCTCCATGCTGACTTGT 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 774 RN.KYH..MH..R.TYSTTDW...HM..S.RY.....N...RCTYT.S.THH..CTYNS 715
QY 287 GTGTATTCACAGTGTGTAACGACCGCTGCTCGGAAACGGTTCTTCTGTGACATTTCTA 346
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 714 BS.B.DNT...H...YNNY.HRNBV.RCA.N.NC..WSCMH.RA.YDD.SNBSBW..T.S 655
QY 347 GAACAGGGCTACAGTCTGATTAATAGATTAATTAAGCCGATTTTGTAGCGCGCGCG 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 654 SBDSYNCB.A...W.RSNN..M.TWANTS.HR..D..A...YN.TAANC..A.B.RCK 595
QY 407 GCGCCGTTTCCCAATAGGAGCGCATTTATGCGCGGAGCTCTACTTTCTTCTTTG 466
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 594 ..AM.HKB..NBRMCNYM.T.S.ANW..HNYTTCMD.A.RH..CY..BSDCK.NT.KY 535
QY 467 GGTAAAGCCCTTCTCTGTTTTCGCGCAGTGTCTGACAGCTGCGCGGAGAACATAGT 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 534 ..TTCCMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KNWC.R.RH 475
QY 527 ATAGAGGATGTAATCTTTCATGTAGAGAAATTAAGCAAGCGGAAAAAACTATGCTACTG 586
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 474 ACB...RDT.C.M.WBH.WWB.AB.HC.W.DG..DB.BKABH..RS.SBS.H...Y..B 415
QY 587 GAGTTGTTTTCATCATATAAAGGAGAAATGTTGCTCACTATGTACAGTTTCTG 646
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 414 ..S.CMN.KTTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHSCYV.TH.YTR 355
QY 647 GAGCTCTTAATCTTTATTTGAGAGAGACTCAATCATACATAGATATTTGCAAAAAA 706
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 ..YBBBCSMABA.YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT. 295
QY 707 AAGACTAATATATAC 721
Db 294 DM.S.M.ATMR...H 280

```

RESULT 15

```

US-10-141-759-198/c
; Sequence 198, Application US/10141759
; Publication No. US20030207361A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C197
; CURRENT APPLICATION NUMBER: US/10/141,759
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 07:32:20 ; Search time 178 Seconds
(without alignments)
6646.225 Million cell updates/sec

Title: US-10-776-213-2

Sequence: 1 cttcgatgacgacgacac.....aaaagacataataacat 723

Scoring table: IDENTITY_NNC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	723	100.0	723	4 US-09-743-194-2	Sequence 2, Appl1
2	723	100.0	11427	4 US-09-743-194-24	Sequence 24, Appl1
3	723	100.0	13073	4 US-09-743-194-20	Sequence 20, Appl1
4	720.4	99.6	850	4 US-09-743-194-30	Sequence 30, Appl1
5	47	6.5	7238	1 US-08-232-463-14	Sequence 14, Appl1
6	37.8	5.2	4989	4 US-09-949-016-17445	Sequence 17445, A
7	37.8	5.2	11440	4 US-09-949-016-17442	Sequence 17442, A
8	36.6	5.1	16662	4 US-09-949-016-16850	Sequence 16850, A
9	35.8	5.0	815	4 US-09-369-247-15	Sequence 15, Appl1
10	35.4	4.9	191433	4 US-09-949-016-16144	Sequence 16144, A
11	35.4	4.9	212449	4 US-09-949-016-15419	Sequence 15419, A
12	35.2	4.9	72992	4 US-09-949-016-17532	Sequence 17532, A
13	34.8	4.8	98962	4 US-09-949-016-14133	Sequence 14133, A
14	34.8	4.8	102884	4 US-09-949-016-17100	Sequence 17100, A
15	34.4	4.8	601	4 US-09-949-016-26160	Sequence 26160, A
16	34.4	4.8	601	4 US-09-949-016-18345	Sequence 18345, A
17	34.4	4.8	51723	4 US-09-949-016-12152	Sequence 12152, A
18	34.4	4.8	51723	4 US-09-949-016-16991	Sequence 16991, A
19	34.4	4.8	105189	4 US-09-949-016-13029	Sequence 13029, A
20	34.2	4.7	1876	3 US-09-469-242-3	Sequence 3, Appl1
21	34.2	4.7	14205	4 US-09-949-016-16196	Sequence 16196, A
22	34.2	4.7	42571	3 US-09-810-347-3	Sequence 3, Appl1
23	34	4.7	57751	4 US-09-949-016-13631	Sequence 13631, A
24	34	4.7	168394	4 US-09-949-016-13002	Sequence 13002, A
25	34	4.7	232547	4 US-09-949-016-16603	Sequence 16603, A
26	33.8	4.7	313	4 US-09-513-999C-35864	Sequence 35864, A
27	33.8	4.7	502	4 US-09-621-976-18891	Sequence 18891, A

28	33.8	4.7	4818	3 US-08-817-926-27	Sequence 27, Appl1
29	33.6	4.6	77851	4 US-09-949-016-12508	Sequence 12508, A
30	33.6	4.6	77867	4 US-09-949-016-13211	Sequence 13211, A
31	33.6	4.6	77867	4 US-09-949-016-13212	Sequence 13212, A
32	33.6	4.6	77940	4 US-09-949-016-12509	Sequence 12509, A
33	33.4	4.6	2172	4 US-09-976-594-125	Sequence 125, App
34	33.4	4.6	312470	4 US-09-949-016-14043	Sequence 14043, A
35	33.4	4.6	336024	4 US-09-949-016-12373	Sequence 12373, A
36	33.2	4.6	784019	4 US-09-949-016-14033	Sequence 14033, A
37	33.2	4.6	828152	4 US-09-949-016-12777	Sequence 12777, A
38	33	4.6	601	4 US-09-949-016-28524	Sequence 28524, A
39	33	4.6	601	4 US-09-949-016-60957	Sequence 60957, A
40	33	4.6	205163	4 US-09-949-016-17009	Sequence 17009, A
41	32.8	4.5	62776	4 US-09-949-016-17576	Sequence 17576, A
42	32.8	4.5	98567	4 US-09-949-016-11750	Sequence 11750, A
43	32.8	4.5	100567	4 US-09-949-016-16934	Sequence 16934, A
44	32.8	4.5	301828	4 US-09-949-016-13969	Sequence 13969, A
45	32.6	4.5	243	4 US-09-248-796A-8463	Sequence 8463, Ap

ALIGNMENTS

RESULT 1					
US-09-743-194-2					
; Sequence 2, Application US/09743194					
; Patent No. 6716601					
; GENERAL INFORMATION:					
; APPLICANT: Bellfield, Graham					
; APPLICANT: Oakley, Caroline					
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for					
; TITLE OF INVENTION: Controlling Nucleic Acid Expression in Yeast					
; FILE REFERENCE: 3526.82543					
; CURRENT APPLICATION NUMBER: US/09/743,194					
; CURRENT FILING DATE: 2001-01-08					
; NUMBER OF SEQ ID NOS: 32					
; SOFTWARE: Patentln Ver. 2.0					
; SEQ ID NO 2					
; LENGTH: 723					
; TYPE: DNA					
; ORGANISM: Saccharomyces cerevisiae					
US-09-743-194-2					
Query Match					
Best Local Similarity 100.0%; Score 723; DB 4; Length 723;					
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	CTTTCGATTAGCAGCAGACATCATAGATGCGTCATATAAATACACTACGAGAA	60		
DB	1	CTTTCGATTAGCAGCAGACATCATAGATGCGTCATATAAATACACTACGAGAA	60		
QY	61	ACCATTAAGACAAAGCATCTACTTTGGAGAGAAAGAGACCGCTTGAAGGGGAT	120		
DB	61	ACCATTAAGACAAAGCATCTACTTTGGAGAGAAAGAGACCGCTTGAAGGGGAT	120		
QY	121	GGGGGCTAAGAGTCACTTCTTTCCCTTCGCGGTCCGAGCCCGGACCCCTCT	180		
DB	121	GGGGGCTAAGAGTCACTTCTTTCCCTTCGCGGTCCGAGCCCGGACCCCTCT	180		
QY	181	CTCCCGCAGATTTCTTCTTCATATCTTCTTTATCTCATCCGTTGAAGAAC	240		
DB	181	CTCCCGCAGATTTCTTCTTCATATCTTCTTTATCTCATCCGTTGAAGAAC	240		
QY	241	GCACCTTAGCTAATGTGTCTGACATCTTCATGCTGTGATCTGTATCTCAGT	300		
DB	241	GCACCTTAGCTAATGTGTCTGACATCTTCATGCTGTGATCTGTATCTCAGT	300		
QY	301	GGTAAAGGACCGTGGCTCGAAACGCTTCTTGTGACAAATTCAGAACAGGGCTACA	360		
DB	301	GGTAAAGGACCGTGGCTCGAAACGCTTCTTGTGACAAATTCAGAACAGGGCTACA	360		
QY	361	GTCTGATATATAGATATATAGGCAATTTTGTCTAGCGCCGCGCGCGCTTCCCA	420		

Db 361 GTCTCGATTAATAGAAATTAATAGCGCAATTTTGTCTAGCGCCGCCGCCGCTTTTCCA 420
Qy 421 ATAGGAGGCGCGATTATTCGCGGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 480
Db 421 ATAGGAGGCGCGATTATTCGCGGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 480
Qy 481 TGTTCCTGCGCAATGTTGCTGCGAGCTGCGCGGAGAACATATGATTAAGGATTTAAC 540
Db 481 TGTTCCTGCGCAATGTTGCTGCGAGCTGCGCGGAGAACATATGATTAAGGATTTAAC 540
Qy 541 TTTCGATGAGAAATTTAGCAAGCGGAAAAAACTATAGGCTAGCTGGAGTGTGTTTCAA 600
Db 541 TTTCGATGAGAAATTTAGCAAGCGGAAAAAACTATAGGCTAGCTGGAGTGTGTTTCAA 600
Qy 601 TCATATTAAGGAGAAATTTGCTCACTATGTGACAGTTTCTGGAGCTTTAACTTT 660
Db 601 TCATATTAAGGAGAAATTTGCTCACTATGTGACAGTTTCTGGAGCTTTAACTTT 660
Qy 661 TATTGCAAGAGACTATCAAAATCTATACAGATATTTGTCAAAAAAAAGACTAATATTA 720
Db 661 TATTGCAAGAGACTATCAAAATCTATACAGATATTTGTCAAAAAAAAGACTAATATTA 720
Qy 721 CAT 723
Db 721 CAT 723

RESULT 2
US-09-743-194-24
; Sequence 24, Application US/09743194
; Patent No. 6716601
; GENERAL INFORMATION:
; APPLICANT: Bellfield, Graham
; APPLICANT: Oakley, Caroline
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/09/743,194
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 11427
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-743-194-24

Query Match 100.0%; Score 723; DB 4; Length 11427;
Best Local Similarity 100.0%; Pred. No. 9.8e-229; Indels 0; Gaps 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTTCGATTAGCAGCAGACATCATGAGCTGCGTATTAATAATACACTACGGAATA 60
Db 15 CTTTCGATTAGCAGCAGACATCATGAGCTGCGTATTAATAATACACTACGGAATA 74
Qy 61 ACCATTAAGAGCAAAAGCATCTTACTTGGAAAGAAAAGAGACAGCTTTGAAGGGGAT 120
Db 75 ACCATTAAGAGCAAAAGCATCTTACTTGGAAAGAAAAGAGACAGCTTTGAAGGGGAT 134
Qy 121 GGGGGCTAAGAGTACTTCTTCTTCCCTTCCGCGTCCGGAACCCGGAACCCCTCT 180
Db 135 GGGGGCTAAGAGTACTTCTTCTTCCCTTCCGCGTCCGGAACCCGGAACCCCTCT 194
Qy 181 CTCCTCCGACGATTTCTTCTTCAATCTTCTTTATTTCTATCCCGTTAAGCAAC 240
Db 195 CTCCTCCGACGATTTCTTCTTCAATCTTCTTTATTTCTATCCCGTTAAGCAAC 254
Qy 241 GCATTAATGATTAATGTGCTGAGACATCTCAATGGCTGTGACTTGTGTATCTCAAGT 300
Db 255 GCATTAATGATTAATGTGCTGAGACATCTCAATGGCTGTGACTTGTGTATCTCAAGT 314
Qy 301 GGTAAAGGACCGTGGCTCGGAACGGTTCTTGTGACAAATTTAGAAAGGGGCTACA 360
Db 361 GGTAAAGGACCGTGGCTCGGAACGGTTCTTGTGACAAATTTAGAAAGGGGCTACA 360

Db 315 GGTAAAGGACCGTGGCTCGGAACGGTTCTTGTGACAAATTTAGAAAGGGGCTACA 374
Qy 361 GTCTCGATTAATAGAAATTAATAGCGCAATTTTGTCTAGCCGCCGCCGCGCTTTTCCA 420
Db 375 GTCTCGATTAATAGAAATTAATAGCGCAATTTTGTCTAGCCGCCGCCGCGCTTTTCCA 434
Qy 421 ATAGGAGGCGCGATTATTCGCGGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 480
Db 435 ATAGGAGGCGCGATTATTCGCGGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 494
Qy 481 TGTTCCTGCGCAATGTTGCTGCGAGCTGCGCGGAGAACATATGATTAAGGATTTAAC 540
Db 495 TGTTCCTGCGCAATGTTGCTGCGAGCTGCGCGGAGAACATATGATTAAGGATTTAAC 554
Qy 541 TTTCGATGAGAAATTTAGCAAGCGGAAAAAACTATAGGCTAGCTGGAGTGTGTTTCAA 600
Db 555 TTTCGATGAGAAATTTAGCAAGCGGAAAAAACTATAGGCTAGCTGGAGTGTGTTTCAA 614
Qy 601 TCATATTAAGGAGAAATTTGCTCACTATGTGACAGTTTCTGGAGCTTTAACTTT 660
Db 615 TCATATTAAGGAGAAATTTGCTCACTATGTGACAGTTTCTGGAGCTTTAACTTT 674
Qy 661 TATTGCAAGAGACTATCAAAATCTATACAGATATTTGTCAAAAAAAAGACTAATATTA 720
Db 675 TATTGCAAGAGACTATCAAAATCTATACAGATATTTGTCAAAAAAAAGACTAATATTA 734
Qy 721 CAT 723
Db 735 CAT 737

RESULT 3
US-09-743-194-20
; Sequence 20, Application US/09743194
; Patent No. 6716601
; GENERAL INFORMATION:
; APPLICANT: Bellfield, Graham
; APPLICANT: Oakley, Caroline
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/09/743,194
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 13073
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-743-194-20

Query Match 100.0%; Score 723; DB 4; Length 13073;
Best Local Similarity 100.0%; Pred. No. 1.1e-228; Indels 0; Gaps 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTTCGATTAGCAGCAGACATCATGAGCTGCGTATTAATAATACACTACGGAATA 60
Db 16 CTTTCGATTAGCAGCAGACATCATGAGCTGCGTATTAATAATACACTACGGAATA 75
Qy 61 ACCATTAAGAGCAAAAGCATCTTACTTGGAAAGAAAAGAGACAGCTTTGAAGGGGAT 120
Db 76 ACCATTAAGAGCAAAAGCATCTTACTTGGAAAGAAAAGAGACAGCTTTGAAGGGGAT 135
Qy 121 GGGGGCTAAGAGTACTTCTTCTTCCCTTCCGCGTCCGGAACCCGGAACCCCTCT 180
Db 136 GGGGGCTAAGAGTACTTCTTCTTCCCTTCCGCGTCCGGAACCCGGAACCCCTCT 195
Qy 181 CTCCTCCGACGATTTCTTCTTCAATCTTCTTTATTTCTATCCCGTTAAGCAAC 240
Db 196 CTCCTCCGACGATTTCTTCTTCAATCTTCTTTATTTCTATCCCGTTAAGCAAC 255
Qy 241 GCATTAATGATTAATGTGCTGAGACATCTCAATGGCTGTGACTTGTGTATCTCAAGT 300
Db 255 GCATTAATGATTAATGTGCTGAGACATCTCAATGGCTGTGACTTGTGTATCTCAAGT 300

Db 256 GCACATGACATAATGCTGACATCTCCATGCTGACCTTGCTGATCTCAAGT 315
QY 301 GGTAACGGACCGCTGCTCGAAACGGTTCTTCTGACAACTCTAAGACAGGGCTACA 360
Db 316 GGTAACGGACCGCTGCTCGAAACGGTTCTTCTGACAACTCTAAGACAGGGCTACA 375
QY 361 GTCTCGATTAATAGATTAATAGCGCAATTTTGTCTAGCGCGCGCGCGCCGTTTCCA 420
Db 376 GTCTCGATTAATAGATTAATAGCGCAATTTTGTCTAGCGCGCGCGCGCCGTTTCCA 435
QY 421 ATAGGAGGCGGAGCTTATGCGGAGCTCTCTTCTTCTTATTTGGGTAAGCCCTTTTC 480
Db 436 ATAGGAGGCGGAGCTTATGCGGAGCTCTCTTCTTCTTATTTGGGTAAGCCCTTTTC 495
QY 481 TGTTCGCGCAGTGTGCTGACAGCTGCGCGGAGAACATAGTATAGGATGTAAC 540
Db 496 TGTTCGCGCAGTGTGCTGACAGCTGCGCGGAGAACATAGTATAGGATGTAAC 555
QY 541 TTTCGATGAGAAATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 600
Db 556 TTTCGATGAGAAATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 615
QY 601 TCATATTAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCTTTAACTTT 660
Db 616 TCATATTAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCTTTAACTTT 675
QY 661 TATTGAGAGGAGCTATCAATCATACAGATTTGTCAAAAAAAGACTAATATATA 720
Db 676 TATTGAGAGGAGCTATCAATCATACAGATTTGTCAAAAAAAGACTAATATATA 735
QY 721 CAT 723
Db 736 CAT 738
RESULT 4
US-09-743-194-30
; Sequence 30, Application US/09743194
; Patent No. 6716601
; GENERAL INFORMATION:
; APPLICANT: Bellfield, Graham
; APPLICANT: Oakley, Caroline
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/09/743,194
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 30
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; US-09-743-194-30
Query Match 99.6%; Score 720.4; DB 4; Length 850;
Best Local Similarity 99.9%; Pred. No.1.5e-228;
Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTTTCGATTAGCAGCAGACACATCAATAGACTGGCTATATAAATAATACATACGAAAA 60
Db 78 CTTTCGATTAGCAGCAGACACATCAATAGACTGGCTATATAAATAATACATACGAAAA 137
QY 61 ACCATTAAGAGCAACCGTACTTCTTGGAGAGAAAGAGAGACCGCTTGTAAAGGGGAT 120
Db 138 ACCATTAAGAGCAACCGTACTTCTTGGAGAGAAAGAGAGACCGCTTGTAAAGGGGAT 197
QY 121 GGGGCTTAAGAACTCACTTTTCCCTTCGCGGTCCGAGCCCGGAGCCCTCCCT 180
Db 198 GGGGCTTAAGAACTCACTTTTCCCTTCGCGGTCCGAGCCCGGAGCCCTCCCT 257
QY 181 CTCGCCGACGATTTCTTCTTCAATATCTTCTTTATCTCTATCCGTTGAAGCAAC 240

Db 258 CTCGCCGACGATTTCTTCTTCAATATCTTCTTTATCTCTATCCGTTGAAGCAAC 317
QY 241 GCACATGACATAATGCTGCTGACATCTCCATGCTGACCTTGCTGATCTCAAGT 300
Db 318 GCACATGACATAATGCTGCTGACATCTCCATGCTGACCTTGCTGATCTCAAGT 377
QY 301 GGTAACGGACCGCTGCTCGAAACGGTTCTTCTGACAACTCTAAGACAGGGCTACA 360
Db 378 GGTAACGGACCGCTGCTCGAAACGGTTCTTCTGACAACTCTAAGACAGGGCTACA 437
QY 361 GTCTCGATTAATAGATTAATAGCGCAATTTTGTCTAGCGCGCGCGCCGTTTCCA 420
Db 438 GTCTCGATTAATAGATTAATAGCGCAATTTTGTCTAGCGCGCGCGCCGTTTCCA 497
QY 421 ATAGGAGGCGGAGTTATGCGGAGCTCTTCTTCTTATTTGGGTAAGCCCTTTTC 480
Db 498 ATAGGAGGCGGAGTTATGCGGAGCTCTTCTTCTTATTTGGGTAAGCCCTTTTC 557
QY 481 TGTTCGCGCAGTGTGCTGACAGCTGCGCGGAGAACATAGTATAGGATGTAAC 540
Db 558 TGTTCGCGCAGTGTGCTGACAGCTGCGCGGAGAACATAGTATAGGATGTAAC 617
QY 541 TTTCGATGAGAAATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 600
Db 618 TTTCGATGAGAAATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 677
QY 601 TCATATTAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCTTTAACTTT 660
Db 678 TCATATTAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCTTTAACTTT 737
QY 661 TATTGAGAGGAGCTATCAATCATACAGATTTGTCAAAAAAAGACTAATATATA 720
Db 738 TATTGAGAGGAGCTATCAATCATACAGATTTGTCAAAAAAAGACTAATATATA 797
QY 721 CA 722
Db 798 AA 799
RESULT 5
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768

;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 16850
;; LENGTH: 16662
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-16850

Query Match 5.1%; Score 36.6; DB 4; Length 16662;
Best Local Similarity 54.0%; Pred. No. 1.1;
Matches 75; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 89 GGAAGAAAAGAGCAGCTGTGAAGGGGATGGGGCTAAGAAGCTTCTTCTT 148
DB 4397 GGAGAAATGAGCAATGATATAAAGGCTCAGCAGCTGCTTTTCTTCTT 4338
QY 149 TCCCTTCGGGTCGGAGCCCGGAGCCCTCTCTCCCGACAGATTCTTTCATAT 208
DB 4337 CTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4278
QY 209 CTTCCTTTATTCCTATCC 227
DB 4277 CTTCCTTCTTCTTCTTCTTCC 4259

RESULT 9
US-09-369-247-15
; Sequence 15, Application US/09369247
; Patent No. 656992
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 44 Human Secreted Proteins
; FILE REFERENCE: P2024PI
; CURRENT APPLICATION NUMBER: US/09/369,247
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/074,118
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,157
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,137
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,341
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,141
; EARLIER FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 815
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (406)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (794)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-369-247-15

Query Match 5.0%; Score 35.8; DB 4; Length 815;
Best Local Similarity 55.1%; Pred. No. 0.34;
Matches 70; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 594 TTTTCATCATATAAAGGAGAAATGTGCTCATATGTGACAGATTCTGGAGAGTCT 653
DB 665 TATCCGTATTTAAATGAGACATTATTTGATCACTGAGTATAGTGTAAATAGT 724
QY 654 TAACTTTATTCGAGAGACTATCAATCATACAGATATTTGCAAAAAAAGACTA 713

DB 725 CTGTGCTATGACAGAGGTGTAATAATTAATATTTTATATAAAAAA 784
QY 714 AATATTA 720
DB 785 AAAAAA 791

RESULT 10
US-09-949-016-16144
; Sequence 16144, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16144
; LENGTH: 191433
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16144

Query Match 4.9%; Score 35.4; DB 4; Length 191433;
Best Local Similarity 59.4%; Pred. No. 12;
Matches 60; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 104 ACGCTTGAAGGGAGATGGGGCTAAGAATCATTCCTTCTTCCCTGCGGTCG 163
DB 41249 ACGTGTGTTAGAGGAGAGAGTGAATCATGTTTCTTCTTCTTCCAGGTAG 41308
QY 164 GACCCGGAGCCCTCTCTCCCGGACGATTTCTTCTTTC 204
DB 41309 TCCCCCCACCCCGGCTGCCCTTCGCGATTTTC 41349

RESULT 11
US-09-949-016-15419/c
; Sequence 15419, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 15419
; LENGTH: 212449
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(212449)
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-15419

Query Match 4.9%; Score 35.4; DB 4; Length 212449;
Best Local Similarity 57.8%; Pred. No. 13;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 123 GGGCTAAGAGTCACTTCTTTTCCCTTGGGGTCCGAGCCCGGACCCCTCTCT 182

DB 124084 GGCTCAATATTCATTCATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 124025

QY 183 CCCCCAGATTTCTCTTCAATCTCTTATTCCTTTCCTTTCCTT 231

DB 124024 CTCCTTCT 123976

RESULT 12

US-09-949-016-17592/c
Sequence 17592, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 17592

LENGTH: 72992

TYPE: DNA

ORGANISM: Human

US-09-949-016-17592

Query Match 4.9%; Score 35.2; DB 4; Length 72992;
Best Local Similarity 60.4%; Pred. No. 7.9;
Matches 58; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 132 AGCTATCACTTCTTCTTCCCTTCCGAGTCCGAGCCCGGACCCCTCTCTCCCGCAG 191

DB 72622 AGCTCTTCT 72563

QY 192 ATTCTCTCTTCAATCTCTCTTATTCCTATCC 227

DB 72562 CTTCT 72527

RESULT 13

US-09-949-016-14133/c
Sequence 14133, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 14133

LENGTH: 98962
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(98962)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14133

Query Match 4.8%; Score 34.8; DB 4; Length 98962;
Best Local Similarity 53.7%; Pred. No. 13;
Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 84 TACTTGAAGAAAGAGCAAGCTTGTAAAGGGAGTGGGGCTTAAGAATTCATCTT 143

DB 78525 TCCTGGAGAGCAAGTGGGTGTGAGTGGGGAGGCTTCAACCAACCCCAACATTT 78466

QY 144 TCTTTTCCCTTCGGGTCCGAGCCCGGACCCCTCTCTCTCCCGAGATTTCTCTT 203

DB 78465 TCCCAACCTTGGCCACACACCCCTGAAACCCCAAGCTCTTAAGCAATATCTTTT 78406

QY 204 CATATCTTCTTTT 217

DB 78405 TTTTCTTTTCTTTT 78392

RESULT 14

US-09-949-016-17100/c
Sequence 17100, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 17100

LENGTH: 102884

TYPE: DNA

ORGANISM: Human

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(102884)

OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17100

Query Match 4.8%; Score 34.8; DB 4; Length 102884;
Best Local Similarity 53.7%; Pred. No. 13;
Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 84 TACTTGAAGAAAGAGCAAGCTTGTAAAGGGAGTGGGGCTTAAGAATTCATCTT 143

DB 78525 TCCTGGAGAGCAAGTGGGTGTGAGTGGGGAGGCTTCAACCAACCCCAACATTT 78466

QY 144 TCTTTTCCCTTCGGGTCCGAGCCCGGACCCCTCTCTCTCCCGAGATTTCTCTT 203

DB 78465 TCCCAACCTTGGCCACACACCCCTGAAACCCCAAGCTCTTAAGCAATATCTTTT 78406

QY 204 CATATCTTCTTTT 217

DB 78405 TTTTCTTTTCTTTT 78392

RESULT 15

US-09-949-016-26160/c
; Sequence 26160, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26160
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-26160

Query Match 4.8%; Score 34.4; DB 4; Length 601;
Best Local Similarity 54.8%; Pred. No. 0.82;
Matches 68; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY	25	CACATGACTGCCTCATTAATACTACGCAAAACCATTAAGACAAAGCATACCT	84
DB	409	CAAAGAGAGAGAGCTCTCAAAACAAACAAACAAACAAACAAACAAACGAGAGAG	350
QY	85	ACTTGAGAGAAAGAGCAGCGCTGTGAAGGGGATGGGCTAAGAGTCATTCACCTT	144
DB	349	ACTGGCCAGAAACAGAGTATATAGCGCAAAAGAGGTGGTGTAGAGAACTTATATATGT	290
QY	145	CTTT	148
DB	289	CATT	286

Search completed: August 27, 2005, 10:04:45
Job time : 182 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 06:52:35 ; Search time 3602 Seconds
(without alignments)
9726.019 Million cell updates/sec

Title: US-10-776-213-2

Perfect score: 723
Sequence: 1 cttcgatagcagcacac.....aaaagactaataaacat 723

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GeneBml:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the local score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	723	100.0	723	6	AR492027 Sequence
2	723	100.0	11427	6	AR492049 Sequence
3	723	100.0	13073	6	AR492045 Sequence
4	720.4	99.6	850	6	AR492055 Sequence
5	720.4	99.6	23498	8	SC9920
6	497.4	68.8	680	6	AX556716 Sequence
7	47	6.5	7218	6	AX556716 Sequence
8	46.8	6.5	257757	3	AE014837
9	43.8	6.1	170295	5	AE014837
10	43.6	6.0	171255	5	AE014837
11	43.6	6.0	208729	2	AC012052
12	42.8	5.9	80659	9	AC004553
13	42.8	5.9	195932	2	AC021710
14	42	5.8	28858	9	AL359542
15	41.8	5.8	209764	10	AL645468
16	41.2	5.7	3291	8	AF210049
17	41	5.7	183978	2	AC024967
18	40.6	5.6	36763	2	AC149338
19	40.4	5.6	248333	2	AC097969

20	40.4	5.6	261990	2	AC11513	AC11513 Rattus no
21	40	5.5	90500	9	AL157889	AL157889 Human DNA
22	40	5.5	139604	9	AC090639	AC090639 Homo sapi
23	40	5.5	208083	2	AC009649	AC009649 Homo sapi
24	40	5.5	227184	2	AC009868	AC009868 Homo sapi
25	40	5.5	238936	2	AC11758	AC11758 Rattus no
26	39.8	5.5	253105	3	PFMAL37	PFMAL37
27	39.4	5.4	152896	2	AC110665	AC110665 Pan trogl
28	39.4	5.4	125723	9	AC099664	AC099664 Homo sapi
29	39.4	5.4	225885	2	AC113236	AC113236 Canis fam
30	39	5.4	2000	6	AX655393	AX655393 Sequence
31	39	5.4	110000	8	CR382131_37	Continuation (38 o
32	39	5.4	150244	9	AC091952	AC091952 Homo sapi
33	39	5.4	168186	2	AC146425	AC146425 Pan trogl
34	39	5.4	174007	2	AC136942	AC136942 Homo sapi
35	38.8	5.4	110000	2	AC117108_0	AC117108 Rattus no
36	38.8	5.4	129547	8	AC006434	AC006434 Genomic s
37	38.8	5.4	137625	9	AC104663	AC104663 Homo sapi
38	38.8	5.4	164310	2	AC016156	AC016156 Homo sapi
39	38.8	5.4	181851	10	AC102413	AC102413 Mus muscu
40	38.8	5.4	203230	2	AC026346	AC026346 Homo sapi
41	38.8	5.4	205592	2	AC150295	AC150295 Callithr
42	38.8	5.4	229634	2	AC150010	AC150010 Callithr
43	38.8	5.4	239450	2	AC106212	AC106212 Rattus no
44	38.6	5.3	2311	6	AX664151	AX664151 Sequence
45	38.6	5.3	100571	9	AL161797	AL161797 Human DNA

ALIGNMENTS

RESULT 1
LOCUS AR492027 723 bp DNA linear PAT 15-MAY-2004
DEFINITION Sequence 2 from patent US 6716601.
ACCESSION AR492027
VERSION AR492027.1 GI:47260496
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 723)
AUTHORS Belfield,G.P. and Oakley,C.
TITLE Compositions and methods utilizing the yeast ZB01 promoter
JOURNAL Patent: US 6716601-A 2 06-APR-2004;
FEATURES
source location/Qualifiers
1..723
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match	100.0%	Score 723;	DB 6;	Length 723;
Best Local Similarity	100.0%	Pred. No. 3.8e-207;	Mismatches 0;	Indels 0;
Matches 723;	Conservative 0;			Gaps 0;
QY	1	CTTTCGATTAGCAGCACACATCATAGATGCGTCATATAAATACACTACGGAAAA	60	
DB	1	CTTTCGATTAGCAGCACACATCATAGATGCGTCATATAAATACACTACGGAAAA	60	
QY	61	ACCATTAAGAGCAAAAGCGATCTTCTTGAAGAAAAGAGCAGCGTTGTAAAGGGGAT	120	
DB	61	ACCATTAAGAGCAAAAGCGATCTTCTTGAAGAAAAGAGCAGCGTTGTAAAGGGGAT	120	
QY	121	GGGGGCTAAGAGTATTCATCTTCTTCCCTGCGGTCGGACCGGGGACCCCTCCT	180	
DB	121	GGGGGCTAAGAGTATTCATCTTCTTCCCTGCGGTCGGACCGGGGACCCCTCCT	180	
QY	181	CTCCCGCAGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	240	
DB	181	CTCCCGCAGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	240	
QY	241	GCACATGACTTAATGCTGGAATCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	300	
DB	241	GCACATGACTTAATGCTGGAATCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	300	

```
Db      241 GCACATATGACATAAGTGTGCTGGACATCTCCATGCTGACTTGTGTATCTCAAGT 300
Qy      301 GGTAACGGACCGCTGGCTCGAAACGGTTCTTTCGTGACAAATTCAGAACAGGGGCTACA 360
Db      301 GGTAACGGACCGCTGGCTCGAAACGGTTCTTTCGTGACAAATTCAGAACAGGGGCTACA 360
Qy      361 GTCCTGATTAATGAATTAATGAAGCGCATTTTGTCTAGCCGCGCGCGCCCGTTTCCCA 420
Db      361 GTCCTGATTAATGAATTAATGAAGCGCATTTTGTCTAGCCGCGCGCGCCCGTTTCCCA 420
Qy      421 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTCTCTATTTTGGGTAAGCCCTTTC 480
Db      421 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTCTCTATTTTGGGTAAGCCCTTTC 480
Qy      481 TGTTCCTGGCCAGTGTGCTGCGAGGCTGCGCGAGAACATATGATTAAGGATGTAAC 540
Db      481 TGTTCCTGGCCAGTGTGCTGCGAGGCTGCGCGAGAACATATGATTAAGGATGTAAC 540
Qy      541 TTTTCGATGAGAAATTTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTCAA 600
Db      541 TTTTCGATGAGAAATTTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTCAA 600
Qy      601 TCATATTAATAAGGAGAAATTTGCTCACTATGTGACATTTCTGGGAGTCTTAACCTT 660
Db      601 TCATATTAATAAGGAGAAATTTGCTCACTATGTGACATTTCTGGGAGTCTTAACCTT 660
Qy      661 TATTGACAGAGACTATCAATCATACAGATATTGTCAAAAAAAAGACTAATATTA 720
Db      661 TATTGACAGAGACTATCAATCATACAGATATTGTCAAAAAAAAGACTAATATTA 720
Qy      721 CAT 723
Db      721 CAT 723

RESULT 2
AR492049      11427 bp      DNA      linear      PAT 15-MAY-2004
LOCUS      AR492049
DEFINITION      Sequence 24 from patent US 6716601.
ACCESSION      AR492049
VERSION      AR492049.1 GI:47260518
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 11427)
AUTHORS      Belfield,G.P. and Oakley,C.
TITLE      Compositions and methods utilizing the yeast ZEO1 promoter
JOURNAL      Patent: US 6716601-A 24 06-APR-2004;
FEATURES
source      Location/Qualifiers
              1..11427
              /organism="unknown"
              /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 723; DB 6; Length 11427;
Best Local Similarity 100.0%; Pred. No. 6.1e-207;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CTTTGATTAGACACGACACATCATAGACTGGGTATTAATAATACACTACGGAATA 60
Db      15 CTTTGATTAGACACGACACATCATAGACTGGGTATTAATAATACACTACGGAATA 74
Qy      61 ACCATTAAGAGCAAGCGATACCTACTTGGAAAGAAAAAGACACGCTTGTAAGGGGAT 120
Db      75 ACCATTAAGAGCAAGCGATACCTACTTGGAAAGAAAAAGACACGCTTGTAAGGGGAT 134
Qy      121 GGGGGCTAAGAAATCATCTTTCTTTCCTTCGCGGTCCGGAACCCGGAACCCCTCCT 180
Db      135 GGGGGCTAAGAAATCATCTTTCTTTCCTTCGCGGTCCGGAACCCGGAACCCCTCCT 194
Qy      181 CTCCTCGACGAAATTCCTTCTTATATCTTCTTTATTCCTATCCGTTGAAGCAAC 240
<
```

```
Db      195 CTCCTCGACGATTTCTTCTTCAATATCTCTTTTATTTCCATATCCGTTGAAGCAAC 254
Qy      241 GCACATATGACATAATGTGTGAGACATCTCCATGCTGACTTGTGTATCTCAAGT 300
Db      255 GCACATATGACATAATGTGTGAGACATCTCCATGCTGACTTGTGTATCTCAAGT 314
Qy      301 GGTAACGGACCGCTGGCTCGAAACGGTTCTTTCGTGACAAATTCAGAACAGGGGCTACA 360
Db      315 GGTAACGGACCGCTGGCTCGAAACGGTTCTTTCGTGACAAATTCAGAACAGGGGCTACA 374
Qy      361 GTCCTGATTAATGAATTAATGAAGCGCATTTTGTCTAGCCGCGCGCGCCCGTTTCCCA 420
Db      375 GTCCTGATTAATGAATTAATGAAGCGCATTTTGTCTAGCCGCGCGCGCCCGTTTCCCA 434
Qy      421 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 480
Db      435 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 494
Qy      481 TGTTCCTGGCCAGTGTGCTGCGAGGCTGCGCGAGAACATATGATTAAGGATGTAAC 540
Db      495 TGTTCCTGGCCAGTGTGCTGCGAGGCTGCGCGAGAACATATGATTAAGGATGTAAC 554
Qy      541 TTTTCGATGAGAAATTTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTCAA 600
Db      555 TTTTCGATGAGAAATTTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTCAA 614
Qy      601 TCATATTAATAAGGAGAAATTTGCTCACTATGTGACATTTCTGGGAGTCTTAACCTT 660
Db      615 TCATATTAATAAGGAGAAATTTGCTCACTATGTGACATTTCTGGGAGTCTTAACCTT 674
Qy      661 TATTGACAGAGACTATCAATCATACAGATATTGTCAAAAAAAAGACTAATATTA 720
Db      675 TATTGACAGAGACTATCAATCATACAGATATTGTCAAAAAAAAGACTAATATTA 734
Qy      721 CAT 723
Db      735 CAT 737

RESULT 3
AR492045      13073 bp      DNA      linear      PAT 15-MAY-2004
LOCUS      AR492045
DEFINITION      Sequence 20 from patent US 6716601.
ACCESSION      AR492045
VERSION      AR492045.1 GI:47260514
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 13073)
AUTHORS      Belfield,G.P. and Oakley,C.
TITLE      Compositions and methods utilizing the yeast ZEO1 promoter
JOURNAL      Patent: US 6716601-A 20 06-APR-2004;
FEATURES
source      Location/Qualifiers
              1..13073
              /organism="unknown"
              /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 723; DB 6; Length 13073;
Best Local Similarity 100.0%; Pred. No. 6.3e-207;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CTTTGATTAGACACGACACATCATAGACTGGGTATTAATAATACACTACGGAATA 60
Db      16 CTTTGATTAGACACGACACATCATAGACTGGGTATTAATAATACACTACGGAATA 75
Qy      61 ACCATTAAGAGCAAGCGATACCTACTTGGAAAGAAAAAGACACGCTTGTAAGGGGAT 120
Db      76 ACCATTAAGAGCAAGCGATACCTACTTGGAAAGAAAAAGACACGCTTGTAAGGGGAT 135
Qy      121 GGGGGCTAAGAAATCATCTTTCTTTCCTTCGCGGTCCGGAACCCGGAACCCCTCCT 180
<
```

Db	136	GGGGGCTAAGAAAGTATTCACTTTCTTTTCCCTTGCGGGTCCGAGCCGCGGACCCCTCCT	195
Qy	181	CTCCCCGACAGATTTCTTCCTTTTCAATACCTTCCTTTATTCCTATCCCGTTGAAGCAACC	240
Db	196	CTCCCCGACAGATTTCTTCCTTTTCAATACCTTCCTTTATTCCTATCCCGTTGAAGCAACC	255
Qy	241	GCACATGACTAAATAGTGTGGAATCATGCCATGGCTGTGACTGTGTGTATCTCAGAGT	300
Db	256	GCACATGACTAAATAGTGTGGAATCATGCCATGGCTGTGACTGTGTGTATCTCAGAGT	315
Qy	301	GGTAAACGGACCGTGGCTCGGAAAACGTTTCCTTCGTGACAAATTCAGAAACAGGGGCTACA	360
Db	316	GGTAAACGGACCGTGGCTCGGAAAACGTTTCCTTCGTGACAAATTCAGAAACAGGGGCTACA	375
Qy	361	GTCCTCGATTAATAGAAATTAATAAGCGCATTTTGTCTAGCGCCGCGCGCGCCCGTTTCCCA	420
Db	376	GTCCTCGATTAATAGAAATTAATAAGCGCATTTTGTCTAGCGCCGCGCGCGCCCGTTTCCCA	435
Qy	421	ATTAGGAGGCGCAGTTTATTCGGCGGAGCTCACTTCCTTACTATTGGGTAAAGCCCTTTC	480
Db	436	ATTAGGAGGCGCAGTTTATTCGGCGGAGCTCACTTCCTTACTATTGGGTAAAGCCCTTTC	495
Qy	481	TGTTTTCGGCGAGTGTGCTGTCAGAGCTCGCGCGGAGAAACAATGATTAAGGATGTAAAC	540
Db	496	TGTTTTCGGCGAGTGTGCTGTCAGAGCTCGCGCGGAGAAACAATGATTAAGGATGTAAAC	555
Qy	541	TTTTGATGATAGGAATTATGACAAACGGAAAAAAACATAATGCTAGCTGGGAGTTGTTTTCAA	600
Db	556	TTTTGATGATAGGAATTATGACAAACGGAAAAAAACATAATGCTAGCTGGGAGTTGTTTTCAA	615
Qy	601	TCATATATAAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCGTCTTAACCTT	660
Db	616	TCATATATAAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCGTCTTAACCTT	675
Qy	661	TATTGACAGAGACTATCAAAATCATACAGATATTGTCAAAAAAAAAAAAAAGACTAATATATA	720
Db	676	TATTGACAGAGACTATCAAAATCATACAGATATTGTCAAAAAAAAAAAAAAGACTAATATATA	735
Qy	721	CAT 723	
Db	736	CAT 738	

[illegible]

Db	138	ACCAATAAAGGCAAGGATATCTACTTTGGAGAGAAAAGAGCAAGCTGTAAAGGGGAT	197
Qy	121	GGGGGGCTAAAGTCAATTCATCTTTCTTTCCCTTCGGGGTCCGAGCCCGGGACCCCTCT	180
Db	198	GGGGGGCTAAAGTCAATTCATCTTTCTTTCCCTTCGGGGTCCGAGCCCGGGACCCCTCT	257
Qy	181	CTCCCGCAGATTTCTTCCCTTCAATATCTTCCCTTATTCCTATCCCGGTGAAGCAAC	240
Db	258	CTCCCGCAGATTTCTTCCCTTCAATATCTTCCCTTATTCCTATCCCGGTGAAGCAAC	317
Qy	241	GCACTATGACTAAATGGTGTGGAACATCTCCATGGCTGTGACTTGTGTATCTCACAGT	300
Db	318	GCACTATGACTAAATGGTGTGGAACATCTCCATGGCTGTGTGTATCTCACAGT	377
Qy	301	GGTAAAGGCAACCGGGCTCGGAAAACGGTCTCTTGTAACAATTTGTAAACAAGGGGCTACA	360
Db	378	GGTAAAGGCAACCGGGCTCGGAAAACGGTCTCTTGTAACAATTTGTAAACAAGGGGCTACA	437
Qy	361	GTCTCGATTAATAGAAATATAAGCGCATTTTGTGTAGCGCGCGCGCCCGCTTTCCCA	420
Db	438	GTCTCGATTAATAGAAATATAAGCGCATTTTGTGTAGCGCGCGCGCCCGCTTTCCCA	497
Qy	421	ATAGGAGGGGCGCATTTATCGGCGGAGCTCTACTTTCTTCTATTTGGGTAAAGCCCTTTC	480
Db	498	ATAGGAGGGGCGCATTTATCGGCGGAGCTCTACTTTCTTCTATTTGGGTAAAGCCCTTTC	557
Qy	481	TGTTTTGGGCAAGTGTGCTGCGAGGCTGCGCGCGAGAACATATGTATAGAGATGTAAAC	540
Db	558	TGTTTTGGGCAAGTGTGCTGCGAGGCTGCGCGCGAGAACATATGTATAGAGATGTAAAC	617
Qy	541	TTTTCGATGAGAAATTAGCAAGCGGAAAAAACTATGCGCTAGCTGGAGTTGTTTTTCAA	600
Db	618	TTTTCGATGAGAAATTAGCAAGCGGAAAAAACTATGCGCTAGCTGGAGTTGTTTTTCAA	677
Qy	601	TCATATATAAGGGGGAATTTGTTGCTACTATGTGACAGTTTCTGGGACGCTTAACTTT	660
Db	678	TCATATATAAGGGGGAATTTGTTGCTACTATGTGACAGTTTCTGGGACGCTTAACTTT	737
Qy	661	TATTCGAGAGACTATCAAAATCATACAGATATTTGTCAAAAAAAGACTAATAATAA	720
Db	738	TATTCGAGAGACTATCAAAATCATACAGATATTTGTCAAAAAAAGACTAATAATAA	797
Qy	721	CA 722	
Db	798	AA 799	

	RESULT	5								
	LOCUS	SC9920								
	DEFINITION	S.cerevisiae chromosome XIII cosmid 9920.								
	VERSION	X48639.1 GI:732924								
	KEYWORDS	CX7; cytochrome oxidase; delta element; glutamate decarboxylae; PETRII; transfer RNA-Ala. <i>Saccharomyces cerevisiae</i> (baker's yeast)								
	SOURCE ORGANISM	<i>Saccharomyces cerevisiae</i> <i>Eukaryota</i> ; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.								
	REFERENCE	1 (bases 1 to 23498) Hunt,S. and Bowman,S. Unpublished								
	AUTHORS	JOURNAL								
	TITLE	Direct Submission								
	JOURNAL	Submitted (10-MAR-1995) <i>Saccharomyces cerevisiae</i> chromosome XIII sequencing project; Sanger Centre, Hinxton Hall, Hinxton, Cambridge CB10 1RQ E-mail: barrell@sanger.ac.uk								
	COMMENT	Notes: All CDS over 100 codons have been analysed. CDS that are completely overlapped and those that are overlapped by more than 50% of their length by a larger CDS have been omitted from this analysis.								

Details of the omitted CDS are available on request. The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons and the calculated codon adaptation index (CAI) is given for each CDS.
Cosmid 9920 is overlapped at the start of this sequence by cosmid 9408

and at the end of this sequence by cosmid 8156.

Location/Qualifiers

1. 23498

FEATURES
source

/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/strain="AB972"

/db_xref="taxon:4932"

/chromosome="XIII"

/map="13R"

/clone="cosmid 9920"

/complement(41..2870)

/note="YM9920.01c, unknown, partial, len: 956, CAI: 0.14;

P500061 Short-chain alcohol dehydrogenase family

signature"

/codon_start=1

/product="unknown"

/protein_id="CA88574.1"

/db_xref="GI:732925"

/db_xref="UniProt/Swiss-Prot:Q04781"

/translation="MSFGINTFOQYNTDPLGSHNGVRIISLNYFDGLPDPISLNLISY

NELKIFSLKREDTEETKEKALMDLSNLISDPNNEYFPENDIFLQMSQIYAKLIISD

YVIRLQSHOITIMLVKSLRKIKSKELKDFILLIGTCELDYSVKPSKINLTCFN

KPKAKINLMAVFOEQLNLVKEIYVNEEDTISDRYSKSESEPRHRYVASAVL

LKLFVHNQVSEKRNSSSLKVLISDESITKLNKKNQONTNAYETVRLIDLVYRGY

MSHKNIMKLAAYKALKLSLTHITSKNILKVCPLPSILMLATLDDYEDGTTWSYDKS

SEKRYLFLVSRTSPSPGFNAVFLVSGTRKSHSLDYLEMLPQWQSGVRLNLEKG

FSARNSAEVLNEFWTNFLFAEDSSEERVKWSEIIFNSLSCGSLSEYTLNTOIS

GVPPDPKRERIEDYTSDEDIRKIVSEKXNLFALLIYSPNNESAISLPPFVOLI

EDPSNVPRKYGVYDALNYFLDSMIPLANGIKIKINIPITLVQESTQNRAGIMAO

YNSKRFKONTDAITSLSDPFIVALSFNLPKTIITATNMLNDIYQULMSDSLE

LVIIDPMKNYKFDSDGEIIFKGNKKFLNQTITTLVRSAVANQVEQFCVLSKDETE

FETLLINDPLFCALYEVSEIDNEKLFKLSIOLAKGNSIANKLAQVILIQHQQVPSF

GAKKRYTAVELINGCNDTSQIFPPANAIEVARYMPAIDYRSSLIVSLSTNTLL

TDKPKYTLKMKOKLIRVALFLDALLDALPERVNNHVARITVSELVTDYNCLSEPN

DIYDPGHTFPRKGRKNLNFSDIVGVNVIOPANGGDMULTFDIAESNSVFFYYSKVL

KVLNLSIDTVSSITTLNGLASVBSFVTKVRQKSTDKOYL"

/complement(2406..2438)

/note="P500061 Short-chain alcohol dehydrogenase family

signature"

/complement(3200..3272)

/product="tRNA-Ala"

/note="tRNA-Ala, anticodon agc, len: 73"

/complement(3379..3706)

/note="delta element"

/complement(4101..4196,4691..4780)

/note="YM9920.02c, unknown, len: 61, CAI: 0.17, possible

small spliced gene"

/codon_start=1

/product="unknown"

/protein_id="CA88575.1"

/db_xref="GI:732926"

/db_xref="UniProt/TrEMBL:O8X1Z0"

/translation="MGEIAPNKKPDEGITNNMCOSLYRYOIMMSIGLRLLPLNKKAI

IFLNSLALHLCALYTE"

/complement(4101..4196,4434..4505)

/note="YM9920.03c, unknown, len: 55, CAI: 0.13, possible

small spliced gene"

/codon_start=1

/product="unknown"

/protein_id="CA88576.1"

/db_xref="GI:732927"

/db_xref="UniProt/TrEMBL:O8X1Y9"

/translation="MELMASCFRMPRYRHANDPTLGLWLTGLRRLPLANKAIIPINSL

ALAHICAIYTE"

/complement(4197..4199)

/note="possible splice acceptor sequence, cag"

misc_feature

CDS

LTR

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

complement(4222..4228)

/note="possible splice branch sequence, tactaac"

/complement(4428..4433)

/note="possible donor sequence, gtaatgt"

/complement(4685..4690)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

/note="possible donor sequence, gtcaagt"

/complement(4689..4694)

CDS
 RSLMTYTLQLYAVFWDVALYKHPNIMSWIGMIIISATLWVIRIRANNETAKDL
 TPIIDDENSILPLEFDLSDSK
 complement (12446..12754)
 /note="Ym9920.08C, unknown, len: 103, CAI: 0.07,
 questionable orf"
 /codon_start=1
 /product="unknown"
 /protein_id="CAA8581.1"
 /db_xref="GI:732932"
 /db_xref="GCA:004838"
 /db_xref="SGD:S0004867"
 /db_xref="UniProt/Swiss-Prot:Q04838"
 /translation="MVPLILILFLFSKFTFLRPVNVHLVTKYTAIVNTKMQTPSII
 DVYTMHVFMYITLILVRKQMSIHAFISGLPSHVLDFSIVRDILSMYLEETAV
 "

CDS
 12832..13398
 /note="Ym9920.09, unknown, len: 188, CAI: 0.15,
 collid=coll domain"
 /codon_start=1
 /product="unknown"
 /protein_id="CAA8582.1"
 /db_xref="GI:732933"
 /db_xref="GCA:004839"
 /db_xref="SGD:S0004868"
 /db_xref="UniProt/Swiss-Prot:Q04839"
 /translation="MPLESIWADAPDEEIPKOKPSHKRSNNKNNKNNRSMNSSEN
 NKKDSVNVKNNKNNHESKTKIKETLIPREKPPHSGKISPVSESLAIDPFGKA
 TEISPPVPSPKMTKTTSKODKPTASKMLKKKIEOREILOKTHHKKQOOVLMDP
 LNDESSNMVDDSEELIQRLKTSIKT"
 complement (13776..13958)

gene

Query Match 99.6%; Score 720.4; DB 8; Length 23498;
 Best Local Similarity 99.9%; Pred. No. 4.2e-206;
 Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTCGATTGACGACACATCATAGATGCGTCATATAAAATACACTACGAAAA 60
 DB 8860 CTTTCGATTGACGACACATCATAGATGCGTCATATAAAATACACTACGAAAA 8919
 QY 61 ACCATAAGAGCAAGGATACCTACTTGGAGAAAAAGAGACCGCTTGAAGGGAGAT 120
 DB 8920 ACCATAAGAGCAAGGATACCTACTTGGAGAAAAAGAGACCGCTTGAAGGGAGAT 8979
 QY 121 GGGGGCTAAGAGTCATCTTCTTTCCCTCGCGGTCGGGACCGCGGACCCCTCCT 180
 DB 8980 GGGGGCTAAGAGTCATCTTCTTTCCCTCGCGGTCGGGACCGCGGACCCCTCCT 9039
 QY 181 CTCGCCGACGATTTCTTCTTTCATATCTTCTTTATTCCTATCCGTTGAAGCAAC 240
 DB 9040 CTCGCCGACGATTTCTTCTTTCATATCTTCTTTATTCCTATCCGTTGAAGCAAC 9099
 QY 241 GCACATAGCTAATGGTGTGACATCTCCATGGCTGTGACTTGTGTATCTCAAGT 300
 DB 9100 GCACATAGCTAATGGTGTGACATCTCCATGGCTGTGACTTGTGTATCTCAAGT 9159
 QY 301 GGTAAAGGACCGTGGTCGGAAAAAGTTCCTTCGTAACAATTCTGAACAGGGGCTACA 360
 DB 9160 GGTAAAGGACCGTGGTCGGAAAAAGTTCCTTCGTAACAATTCTGAACAGGGGCTACA 9219
 QY 361 GTCTCGATAATATAAGGCAATTTTGTAGAGCGCGCGGCGCGCCCTTTCCCA 420
 DB 9220 GTCTCGATAATATAAGGCAATTTTGTAGAGCGCGCGCGCGCCCTTTCCCA 9279
 QY 421 ATAGGAGGCGAGTTTATCGGCGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTTC 480
 DB 9280 ATAGGAGGCGAGTTTATCGGCGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTTC 9339
 QY 481 TGTTCGCGCAGGTGTGTGAGAGCTGGCGGAGAAATATGATGTAAGGATGTAAC 540
 DB 9340 TGTTCGCGCAGGTGTGTGAGAGCTGGCGGAGAAATATGATGTAAGGATGTAAC 9399
 QY 541 TTTGATGAGAGATTAGCAAGCGGAAAAAACTATGAGCTAGCTGGAGTTGTTTTTCAA 600

DB 9400 TTTGATGAGAGATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 9459
 QY 601 TCATATTAAGAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCTTAACTTT 660
 DB 9460 TCATATTAAGAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCTTAACTTT 9519
 QY 661 TATTCGAGAGACTATCAATCATACAGATATTGTTCAAAAAAGACTTAATATA 720
 DB 9520 TATTCGAGAGACTATCAATCATACAGATATTGTTCAAAAAAGACTTAATATA 9579
 QY 721 CA 722
 DB 9580 AA 9581

RESULT 6
 AX536716 680 bp DNA linear PAT 22-NOV-2002
 LOCUS
 DEFINITION Sequence 317 from Patent WO02064766.
 AX536716
 ACCESSION
 VERSION AX536716.1 GI:25263152
 KEYWORDS
 SOURCE Saccharomyces cerevisiae (baker's yeast)
 ORGANISM Saccharomyces cerevisiae
 Buxarjola; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE
 1 Contreras, R.H., Eberhardt, I., Luyten, W.H. and Reekmans, R.J.
 Bax-responsive genes for drug target identification in yeast and
 fungi
 Patent: WO 02064766-A 317 22-AUG-2002;
 JOURNAL PHARMACEUTICA N.V. (BE)
 FEATURES
 source location/Qualifiers
 1..680
 /organism="Saccharomyces cerevisiae"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4932"

ORIGIN

Query Match 68.8%; Score 497.4; DB 6; Length 680;
 Best Local Similarity 99.8%; Pred. No. 6.7e-139;
 Matches 498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 224 ATCCGTTGAAGCAACCGCACTATGACTTAATGTGTCTGACATCTCATGCTGTGACT 283
 DB 1 ATCCGTTGAAGCAACCGCACTATGACTTAATGTGTCTGACATCTCATGCTGTGACT 60
 QY 284 TGTGTATCTCAAGTGTGAACGGACCGTGGCTCGAAACGGTTCTTGTGACAAAT 343
 DB 61 TGTGTATCTCAAGTGTGAACGGACCGTGGCTCGAAACGGTTCTTGTGACAAAT 120
 QY 344 CTAGAACAGGGGCTACAGTCTCGATATATGAATTAATAGCGCAATTTTCTAGCGCGCC 403
 DB 121 CTAGAACAGGGGCTACAGTCTCGATATATGAATTAATAGCGCAATTTTCTAGCGCGCC 180
 QY 404 GGGGGCCCGTTTCCCAATAGGAGGCGAGTTTATCGCGGAGCTCTACTTCTTCTAT 463
 DB 181 GGGGGCCCGTTTCCCAATAGGAGGCGAGTTTATCGCGGAGCTCTACTTCTTCTAT 240
 QY 464 TTGGGTAAGCCCTTTCTGTGTTTCGCGCAGTGTGCTGACAGCTGCGCGGAGAACATA 523
 DB 241 TTGGGTAAGCCCTTTCTGTGTTTCGCGCAGTGTGCTGACAGCTGCGCGGAGAACATA 300
 QY 524 GTGATAGGAGATGAATCTTTCGATGAGAGATTTAGCAAGCGGAAAAAACTATGGCTAC 583
 DB 301 GTGATAGGAGATGAATCTTTCGATGAGAGATTTAGCAAGCGGAAAAAACTATGGCTAC 360
 QY 584 TGGGAGTTGTTTTTAATCATATAAAGGAGAAATTTGTGCTACTATGTGACAGTTTC 643
 DB 361 TGGGAGTTGTTTTTAATCATATAAAGGAGAAATTTGTGCTACTATGTGACAGTTTC 420
 QY 644 TGGGAGTTTAACTTTATTCAGAGGACTATCAATCATACAGATATTGTCAAAAAA 703

Db 421 TGGGAGCTCTTAATTGATGAGAGACTATCAATCATACGATATTGTCAAAAAA 480
Qy 704 AAAAAAGCTAATAATAACA 722
Db 481 AAAAAAGCTAATAATAAAA 499

RESULT 7
LOCUS 166494 7218 bp DNA linear PAT 28-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION 166494
VERSION 166494.1 GI:2724471
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7218)
AUTHORS Dörner, F., Scheiflinger, F. and Falkner, F. Gunter.
JOURNAL Recombinant fowlpox virus
PATENT: US 5670367-A 14 23-SEP-1997;
FEATURES
source location/Qualifiers
1..7218
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 6.5%; Score 47; DB 6; Length 7218;
Best Local Similarity 2.9%; Pred. No. 0.019;
Matches 11; Conservative 212; Mismatches 152; Indels 0; Gaps 0;

Qy 113 AGGGGAGATGGGGGCTAAGAACTATTCCTTTCCCTTGGCGGACCGGGA 172
Db 1047 AGGTGAGGAGGAGCTTCGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1106
Qy 173 CCCCTCCTCCCGGACGATTTCTCTTCAATCTCTTATCTTATTCGATCCGCTG 232
Db 1107 TT 1166
Qy 233 AAGCAACCGCATGATCAATTAATGAGTGCAGCATCTCATGCTGATGTTGTTAT 292
Db 1167 TT 1226
Qy 293 CTCACAGTGTAAAGGACCGTGGCTCGAAGCGTTCCTTGCTGCAATCTAGAAG 352
Db 1227 TT 1286
Qy 353 GGGCTACAGTCTCGAATATAGATTAATAGCGATTTTGCATAGCCGCCCGCGGCC 412
Db 1287 TT 1346
Qy 413 GTTCCCAATAGGAGCGCATTTATCGGCGAGCTTACTTCTTCTATTGGGTAG 472
Db 1347 TT 1406
Qy 473 CCCCTTCTGTTTC 487
Db 1407 TTTTTTTTTTTTTT 1421

RESULT 8
LOCUS AE014837 257757 bp DNA linear INV 07-OCT-2002
DEFINITION Plasmodium falciparum 3D7 chromosome 11 section 2 of 8 of the
complete sequence.
ACCESSION AE014837 AE014186
VERSION AE014837.1 GI:2349596
KEYWORDS
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
REFERENCE 1 (bases 1 to 257757)
AUTHORS Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W.,

Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulsen, I. T.,
James, K., Eisen, J. A., Rutherford, K., Salzberg, S. L., Craig, A.,
Kyes, S., Chan, M. S., Nene, V., Shallow, S. J., Sub, B., Peterson, J.,
Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D.,
Mather, M. W., Vaidya, A. B., Martin, D. M. A., Fairclamb, A. H.,
Fraunholz, M. J., Roos, D. S., Ralph, S. A., McFadden, G. I.,
Cummings, L. M., Subramanian, G. M., Mungall, C., Venter, J. C.,
Carucci, D. J., Hoffman, S. L., Newbold, C., Davatz, R. W., Fraser, C. M. and
Barrell, B.
TITLE Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL Nature 419 (6906), 498-511 (2002)
PUBMED 12368864
REFERENCE 2 (bases 1 to 257757)
AUTHORS Gardner, M. J.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
FEATURES
source location/Qualifiers
1..257757
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="11"
repeat_region 257..381
/rpt_type=tandem
repeat_region 305..373
/rpt_type=tandem
repeat_region 763..787
/rpt_type=tandem
repeat_region 1304..1333
/rpt_type=tandem
gene complement(1434..1828,1910..1994)
/locus_tag="PF11_0070"
/locus_tag="pf11_0070"
complement(join(1434..1828,1910..1994))
/locus_tag="pf11_0070"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA035659.1"
/db_xref="GI:23495997"
/translation="MCMVCIIECTEKKKKKKKKKLPIYIHKIDVKKAEONPL
VPRYIGHLFEKROFSTLYFVTTINMDVKIHAVALVAERRSLEFVGIKK
ANIYIVNKKRKPKEIPNISKISPSIAVINSKDEITPSKVLFLFISLSTKIV"
1872..1902
repeat_region 1926..1960
/rpt_type=tandem
repeat_region 2199..2240
/rpt_type=tandem
repeat_region 2384..2542
/rpt_type=tandem
repeat_region 2582..2650
/rpt_type=tandem
gene complement(2647..2650)
/locus_tag="PF11_0071"
/locus_tag="pf11_0071"
complement(2647..2650)
/locus_tag="PF11_0071"
/codon_start=1
/product="PuvB DNA helicase, putative"
/protein_id="AA035660.1"
/db_xref="GI:23495998"
/translation="MEINISNITSSNNKKERNVNIHSHIKIGVNTNIVLHENSIS
DEKTSMPFDNSGCVGGFKAREASLFLVDLIKQKKAAGKCTLLKPGSGSGSALAIGT
SREINRMPFVFLSGSEVSNIEIKRTVLLEAFKRSJHIKIKERLVBEGVDMVVE
ENECIYSLNKAQJINAIITLKSVMGSKTULAPRKHIOIAREKIKIGDVIYITNG
HKKRLGNDVAKEVDIEFDEYSLPKGEVAKKEVVOQSILHIDLANAPVIGEDL
ASVNVYLRPKTEITETKRYEINKTVNKLFSGLAFIIEGVLIIDEAHMDIGCESY
LNRAIESPLAIIIMATNRGICVYKGDNIENPHSIPVDLDRLIITTFEPTLAEIV
ITLARAQTEKINITEEGNNYIAKIGIOSLRPFAMLLPESRIITLBSNTIIDIKHI

repeat__region	OADELFMADKTSARHVVDQNKNEVN"	4159. .4258
repeat__region	/rpt_type=candem	
repeat__region	4159. .4215	
repeat__region	/rpt_type=candem	
repeat__region	4718. .4749	
repeat__region	/rpt_type=candem	
repeat__region	4922.4978	
repeat__region	/rpt_type=candem	
repeat__region	5175. .5199	
repeat__region	/rpt_type=candem	
repeat__region	5322. .5352	
repeat__region	/rpt_type=candem	
repeat__region	5350. .5397	
repeat__region	/rpt_type=candem	
repeat__region	5405. .5438	
repeat__region	/rpt_type=candem	
repeat__region	5531. .5557	
repeat__region	/rpt_type=candem	
gene	<5712. .>6914	
mRNA	/locus_tag="PF11_0072"	
CDS	join(<5712. .5802,5896. .6191,6311. .6454,6636. .>6914)	
	/locus_tag="PF11_0072"	
	join(5712. .5802,5896. .6191,6311. .6454,6636. .6914)	
	/locus_tag="PF11_0072"	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="AAN35661.1"	
	/db_xref="GI:23495999"	
repeat__region	/translation="MKNIKSCVCLLPFYIOIWINCLPFGYNIKNLGVPELTIK	
repeat__region	KNVQGRKQOLIKVQEGKENIPTDGDNTNTIKEDTKQRTNIDIKVPELTIK	
repeat__region	NEILKSENIULERSFSEKEDITTTIPELLKLTIOYKRKIFQRDKCGSSLEIOIIL	
repeat__region	LTPEILPFLTEMKKKKKDPAFLGRFLPCVSKRRRLVYLGRDKREMEKTIYSYNIKK	
repeat__region	PLLRTEPEYVKDLYIHFNVTGRFKNAEKKKKKKLKKKKVQTDKILFGN"	
repeat__region	6250. .6288	
repeat__region	/rpt_type=candem	
repeat__region	6482. .6530	
repeat__region	/rpt_type=candem	
repeat__region	6484. .6533	
repeat__region	/rpt_type=candem	
repeat__region	6913. .6938	
repeat__region	/rpt_type=candem	
repeat__region	7132. .7156	
repeat__region	/rpt_type=candem	
repeat__region	7287. .7327	
repeat__region	/rpt_type=candem	
repeat__region	7691. .7726	
gene	/rpt_type=candem	
gene	<7800. .>10244	
mRNA	/locus_tag="PF11_0073"	
CDS	join(<7800. .9712,9978. .10049,10155. .10162,10204. .>10244)	
	/locus_tag="PF11_0073"	
	join(7800. .9712,9978. .10049,10155. .10162,10204. .10244)	
	/locus_tag="PF11_0073"	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="AAN35662.1"	
	/db_xref="GI:23496000"	
	/translation="MPCSEFLPFPFPVQNKELAILSKVHLKHHIYPRKSHFPI	
	YSDVNNVNNLKNKETSJKDKRIKYEDVHNNKKKIFLPFLFLKNNNDVYITEBH	
	KDKNIIDNNNNKKDKDDIINNDDKKNIINDKQNTINDKQNSNFFETENNLLK	
	LKNNLEVFSSYINXSIDIIVPLNYHDLNKLHLIEIIPNNIINSLDYIFLNKKDYVK	
	SDIINNKNNNNNNNKYNICPEYNNIYNIIFPIYVNDQNPINISQYTSFPTNKK	
	DNIIHKKPOLVHVIQKFLNDHYNLKDQREHKYININPFLTKNIDILINDKDIEE	
	NKMLYVKKKKILKKEYNVNTNTNTNTNTRKQEDQULIYVHKNDTPFKKKYKPPRLKN	
	LTNENYITNPSSYKLDIKETIYAKOITTPNNNVIKHIONIKKKHNIIEILN	
	NFKKKKNNKINSVHNVQETIKKKQYLYLHIIISLDLEKYSNLEIKNLYIYQFKEN	
	IQKIKNSDILPQKKIINHIDQLPQYNSKFIIPSNKNYQAQGSYYIYYKXIKL	
	LQNLQANELINIVYISKGLYVKNYSFNFSKSYTILNIYITIFLILKNNKNIWF	
	NYSLPNAPGSSYXDDISLPKKDIMITTSEMDQYQKIYIQISLVVTPPKKCS"	
repeat__region	8865. .8890	
repeat__region	/rpt_type=candem	
repeat__region	9936. .9975	

[illegible]

SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 170295)
 AUTHORS Hunter, G.
 TITLE Direct Submission
 JOURNAL Submitted (06-JAN-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk
 COMMENT On Jan 6, 2004 this sequence version replaced gi:37805672.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.
 Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'dir' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmaek.shtml
 is from a CHORI-211 BAC library
 VECTOR: pTARBAC2.1.
 Location/Qualifiers
 1..170295
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="CH211-144B6"
 /clone_1ib="CHORI-211"

FEATURES
 source
 1..170295
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="CH211-144B6"
 /clone_1ib="CHORI-211"

ORIGIN

Query Match 6.1%; Score 43.8; DB 5; Length 170295;
 Best Local Similarity 57.8%; Pred. No. 0.3; Mismatches 0; Gaps 0;
 Matches 78; Conservative 0; Indels 57; Indels 0; Gaps 0;
 588 AGTGTGTTTCAATCAATATAAAGGAGAAATGTTGCTACTATGTGACAGTTTCTGGG 647
 60267 ATTATATATCAATATATAAATATATATATATATATATATATATATATATATG 60326
 648 AGCTCTACTTTTATGAGAGACTATCAATCTACAGATATTTGTAATAAAAAA 707
 60327 TTGTATATATATATATAAACAACAACAACATATAAATAATTAACAAATATATA 60386
 708 AGACTAATATATAACA 722
 60387 ATACTAATATAATA 60401

RESULT 10
 AC092573/c 171265 bp DNA linear PRI 01-MAR-2002
 LOCUS Homo sapiens BAC clone RP11-107 from 2, complete sequence.
 DEFINITION AC092573 AC015764
 AC092573.2 GI:15668084
 VERSION AC092573.2 GI:15668084
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 171265)
 Sulston, J.E. and Waterston, R.
 Toward a complete human genome sequence
 JOURNAL Medical Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 2 (bases 1 to 171265)
 Tomlinson, C., Cotton, M., Elliott, G., Dixon, R., Hawkins, M. and Boyer, E.
 The sequence of Homo sapiens BAC clone RP11-107
 JOURNAL Unpublished (2001)
 3 (bases 1 to 171265)
 Waterston, R.H.
 Direct Submission
 Submitted (19-JUN-2001) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108 USA
 4 (bases 1 to 171265)
 Waterston, R.H.
 Direct Submission
 Submitted (19-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108 USA
 5 (bases 1 to 171265)
 Waterston, R.
 Direct Submission
 Submitted (01-MAR-2002) Department of Genetics, Washington University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Sep 19, 2001 this sequence version replaced gi:14916158.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0001007
 Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Moon, P.Y., Zhao, B., Frengen, E.,

Qy 721 CA 722
Db 55928 CA 55927

RESULT 11
AC012052/c
LOCUS
DEFINITION
AC012052 208729 bp DNA linear HTG 04-MAY-2001
Homo sapiens chromosome 2 clone RP11-337F19 map 2, WORKING DRAFT
SEQUENCE, 23 unordered pieces.
AC012052
AC012052.3 GI:8576227
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 2, clone RP11-337F19
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 208729)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Balwin,J., Barna,N., Beckerly,R., Boguslavsky,I., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domingo,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heathford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatsis,A., Klein,D.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Margulis,N.,
McEwan,P., McGuirk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tittrell,A., Vassiliev,H., Vo,A., Wheeler,D., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
JOURNAL Direct Submission
COMMENT Submitted (19-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 21, 2000 this sequence version replaced gi:7321517.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center Project name: L1385
Center clone name: 337_F_19

Summary Statistics
Sequencing vector: M13; M77815, 100% of reads
Chemistry: Dye-primer-amersham; 3% of reads
Chemistry: Dye-terminator Big Dye; 97% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 191830 bases at least Q40
Consensus quality: 199935 bases at least Q30
Consensus quality: 203385 bases at least Q20
Insert size: 210000; agarose-fp
Insert size: 206529; sum-of-ctrls
Quality coverage: 4.0 in Q20 bases; agarose-fp
Quality coverage: 4.1 in Q20 bas.

NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1080: contig of 1080 bp in length

1081 1180: gap of 100 bp
* 1181 2268: contig of 1088 bp in length
* 2269 2368: gap of 100 bp
* 2369 3746: contig of 1378 bp in length
* 3747 3846: gap of 100 bp
* 3847 6128: contig of 2282 bp in length
* 6129 6428: gap of 100 bp
* 6429 7907: contig of 1679 bp in length
* 7908 8007: gap of 100 bp
* 8008 10148: contig of 2141 bp in length
* 10149 10248: gap of 100 bp
* 10249 12611: contig of 2363 bp in length
* 12612 12711: gap of 100 bp
* 12712 14921: contig of 2210 bp in length
* 14922 15021: gap of 100 bp
* 15022 18148: contig of 3127 bp in length
* 18149 18248: gap of 100 bp
* 18249 22510: contig of 4262 bp in length
* 22511 22610: gap of 100 bp
* 22611 27371: contig of 4761 bp in length
* 27372 27471: gap of 100 bp
* 27472 32405: contig of 4934 bp in length
* 32406 32505: gap of 100 bp
* 32506 38167: contig of 5662 bp in length
* 38168 38267: gap of 100 bp
* 38268 46372: contig of 8105 bp in length
* 46373 46472: gap of 100 bp
* 46473 56551: contig of 10079 bp in length
* 56552 56651: gap of 100 bp
* 56652 64715: contig of 8064 bp in length
* 64716 64815: gap of 100 bp
* 64816 75459: contig of 10644 bp in length
* 75460 75559: gap of 100 bp
* 75560 90744: contig of 15185 bp in length
* 90745 90844: gap of 100 bp
* 90845 107238: contig of 16394 bp in length
* 107239 107338: gap of 100 bp
* 107339 128875: contig of 21537 bp in length
* 128876 128975: gap of 100 bp
* 128976 152422: contig of 23447 bp in length
* 152423 152522: gap of 100 bp
* 152523 178409: contig of 25887 bp in length
* 178410 178510 208729: gap of 100 bp
* 208729: contig of 30220 bp in length.

FEATURES
source
1..208729
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-337F19"
/clone_lib="RPC1-11 Human Male BAC"
1..1080
/note="assembly_fragment"
1181..2268
/note="assembly_fragment"
clone end:17
vector side:right"
2369..3746
/note="assembly_fragment"
3847..6128
/note="assembly_fragment"
6229..7907
/note="assembly_fragment"
8008..10148
/note="assembly_fragment"
10249..12611
/note="assembly_fragment"
12712..14921
/note="assembly_fragment"
15022..18148
/note="assembly_fragment"
18149..22510
/note="assembly_fragment"

```

misc_feature      18249..22510
                  /note="assembly_fragment"
misc_feature      22611..27371
                  /note="assembly_fragment"
misc_feature      27472..32405
                  /note="assembly_fragment"
misc_feature      32506..38167
                  /note="assembly_fragment"
misc_feature      38268..46372
                  /note="assembly_fragment"
misc_feature      46473..56551
                  /note="assembly_fragment"
misc_feature      56652..64715
                  /note="assembly_fragment"
misc_feature      64816..75459
                  /note="assembly_fragment"
misc_feature      75560..90744
                  /note="assembly_fragment"
misc_feature      90845..107238
                  /note="assembly_fragment"
misc_feature      107339..128875
                  /note="assembly_fragment"
misc_feature      128976..152422
                  /note="assembly_fragment"
                  clone_end:SP6
                  vector_side:left"
misc_feature      152523..178409
                  /note="assembly_fragment"
misc_feature      178510..208729
                  /note="assembly_fragment"

```

ORIGIN

Query Match 6.0%; Score 43.6; DB 2; Length 208729;

```

Best Local Similarity 56.0%; Pred.No.0.35; Indels 1; Gaps 1;
Matches 102; Conservative 0; Mismatches 79;
QY      542 TTGAGAGAGAGATTAGCAAGCGAATAAATCATGCTAGCTGGAGTTGTTTTCAT 601
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      73634 TTTAGTAAAAAATAAATAAATAAATAAATGCTGGTTTGTTACTCTGTTTGA 73575
QY      602 CATATAAAAAGGAGAAATTGTTGCTCAGTATGTGACAGTTTCTGGAGCTTTA-AC 660
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      73574 GCTATAAAAAGGAGAAATTATATCAACATCTCACACATTCGGAGGAAAAATAA 73515
QY      661 TATTCAGAGAGATATCAATCTATACAGATATGTCAAAAAAATAAGACTAATAATA 720
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      73514 TCTGAGCTAGGCTCTCAAGGCTCTTAGAGACGAGAGAGAAACAAAGATGCA 73455
QY      721 CA 722
        ||
Db      73454 CA 73453

```

```

RESULT 12
AC004553/c      80659 bp      DNA      linear      PRI 25-MAY-2002
LOCUS           Homo sapiens X GSHB-223P11 (Genome Systems Human BAC library)
DEFINITION      Complete sequence.
ACCESSION       AC004553
VERSION         AC004553.1 GI:3451332
KEYWORDS        HTG.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS         Muzny,D., Aronson,A.D., Adams,C., Brundage,E., Bunce,C.,
                Carvelli,K., Chacko,J., Chen,J., Di,W., Ding,Y., Dugan,S.,
                Durbin,J., Forcum,J., Ganesh,R., Garcia,C., Goodman,M.,
                Gorrell,J.H., Haywood,M., Hernandez,J., Jackson,L., Jin,S.,
                Kampel,R., Karpaty,S., Kovar,C., Leal,B., Li,Y., Licharge,O.,
                Liu,W., Logan,O., Lu,J., Martinez,C., Oswal,G., Perez,L.,
                Rashid,N.D., Rowland,K., Savage,L., Scherer,S.E., Shen,H.,

```

```

TITLE           Simon,M., Stoval,K., Timms,K.M., Todd,J., Vo,Q., Williamson,A.,
                Worley,K.C., Yu,W., Chinnault,C., Nelson,D. and Gibbs,R.A.
JOURNAL         Unpublished
REFERENCE       2 (bases 1 to 80659)
AUTHORS         Worley,K.C.
TITLE           Submitted (11-APR-1998) Molecular and Human Genetics, Baylor
                College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE       3 (bases 1 to 80659)
AUTHORS         Worley,K.C.
TITLE           Direct Submission
                Submitted (25-AUG-1998) Human Genome Sequencing Center, Department
                of Molecular and Human Genetics, Baylor College of Medicine, One
                Baylor Plaza, Houston, TX 77030, USA
REFERENCE       4 (bases 1 to 80659)
AUTHORS         Worley,K.C.
TITLE           Direct Submission
                Submitted (17-SEP-1998) Human Genome Sequencing Center, Department
                of Molecular and Human Genetics, Baylor College of Medicine, One
                Baylor Plaza, Houston, TX 77030, USA
REFERENCE       5 (bases 1 to 80659)
AUTHORS         Worley,K.C.
TITLE           Direct Submission
                Submitted (29-MAR-2002) Human Genome Sequencing Center, Department
                of Molecular and Human Genetics, Baylor College of Medicine, One
                Baylor Plaza, Houston, TX 77030, USA
REFERENCE       6 (bases 1 to 80659)
AUTHORS         Worley,K.C.
TITLE           Direct Submission
                Submitted (25-MAY-2002) Human Genome Sequencing Center, Department
                of Molecular and Human Genetics, Baylor College of Medicine, One
                Baylor Plaza, Houston, TX 77030, USA

```

COMMENT

On Aug 25, 1998 this sequence version replaced gi:3402643. Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

```

FEATURES
source          1..80659
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="X"
                /clone="GSHB-223P11"
                1..1999
                /note="overlaps bases 114570..116568 of clone AC112492"
                /function="clone overlap"
                1310..1350
                /rpt_family="L2"
                1527..1639
                /rpt_family="MTR"
                /complement(1641..1930)
                /rpt_family="AluY"
                1933..2050
                /rpt_family="MIR"
                /complement(3587..3670)
                /rpt_family="MIR"
                3679..3861
                /rpt_family="MER53"
                4687..4829
                /rpt_family="MER5B"

```

```
repeat_region complement(5319,.5358)
/rpt_family="AT_rich"
repeat_region complement(5853,.5898)
/rpt_family="AT_rich"
repeat_region complement(5943,.6244)
/rpt_family="MLT1B"
repeat_region complement(6228,.6290)
/rpt_family="MLT1D"
repeat_region complement(6354,.6624)
/rpt_family="L1MA2"
repeat_region complement(6631,.6758)
/rpt_family="MLT1D"
repeat_region complement(7635,.7690)
/rpt_family="(CA)n"
repeat_region 7833,.8023
/rpt_family="MER3"
repeat_region 8188,.8352
/rpt_family="MER5A"
repeat_region 9035,.9057
/rpt_family="AT_rich"
repeat_region 9148,.9227
/rpt_family="MIR"
repeat_region complement(9246,.9302)
/rpt_family="AT_rich"
repeat_region complement(9663,.10244)
/rpt_family="L1MC4"
repeat_region complement(10630,.10701)
/rpt_family="(TA)n"
repeat_region complement(11161,.11220)
/rpt_family="(GA)n"
repeat_region complement(12725,.12846)
/rpt_family="(TA)n"
repeat_region complement(12853,.12932)
/rpt_family="(CATA)n"
repeat_region 13300,.13428
/rpt_family="MIR"
repeat_region complement(13783,.14150)
/rpt_family="L1PA8"
repeat_region 14162,.14188
/rpt_family="AT_rich"
repeat_region complement(14686,.14806)
/rpt_family="FLAM_C"
repeat_region 15736,.15925
/rpt_family="L2"
repeat_region complement(16745,.17110)
/rpt_family="L1M4"
repeat_region 17480,.17941
/rpt_family="MLT1C"
repeat_region 18083,.18134
/rpt_family="MADE1"
repeat_region complement(18177,.18784)
/rpt_family="L1MA9"
repeat_region 19560,.19858
/rpt_family="AluSg"
repeat_region complement(20203,.20229)
/rpt_family="AT_rich"
repeat_region 20350,.20551
/rpt_family="MER4_internal"
repeat_region 20563,.21094
/rpt_family="MER4_internal"
repeat_region 21096,.21246
/rpt_family="MER4A2"
repeat_region 21247,.21510
/rpt_family="MER4A"
repeat_region 21804,.22111
/rpt_family="AluSp"
repeat_region complement(22905,.22958)
/rpt_family="MIR"
repeat_region 23531,.24248
/rpt_family="L1ME3A"
repeat_region complement(24998,.25210)
/rpt_family="L1PA8"
repeat_region 25211,.25359
```

```
repeat_region /rpt_family="MER11A"
25357,.26294
/rpt_family="MER11B"
repeat_region complement(26295,.26395)
/rpt_family="L1PA8"
repeat_region 26608,.26927
/rpt_family="AluSg"
repeat_region 27578,.27878
/rpt_family="AluSx"
repeat_region complement(28196,.28279)
/rpt_family="L2"
repeat_region complement(28709,.28748)
/rpt_family="(TAA)n"
repeat_region complement(28749,.29386)
/rpt_family="L1PA4"
STS 29091,.29222
/standard_name="G19948"
/db_xref="dbSTS:32826"
repeat_region 29361,.29554
/rpt_family="L1P"
repeat_region complement(30273,.31278)
/rpt_family="MER11A"
repeat_region complement(32126,.32237)
/rpt_family="(GAA)n"
STS 32378,.32676
/standard_name="DXS67"
complement(33176,.33223)
/rpt_family="MIR"
repeat_region 33928,.34051
/rpt_family="MIR"
repeat_region complement(35042,.35071)
/rpt_family="AT_rich"
repeat_region complement(35418,.35439)
/rpt_family="AT_rich"
repeat_region complement(38090,.38123)
/rpt_family="AT_rich"
```

Query Match 5.9%; Score 42.8; DB 9; length 80659;
Best Local Similarity 51.0%; Pred. No. 0.52;
Matches 101; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

```
QY 517 GAACATAGTGAATAGGATGTAACCTTTCGATGAGAGAAATTCGACAGCGGAAAAACTAT 576
|||||
Db 33326 GAACATAGTTCGTCAGAGCTTCACAACTCCGAGTTCAAAGTTCAGATGTTTAAAGAAATPA 33267

QY 577 GGCTACGCTGGAGGAGTGTTCATCATATATAAAGGAGAAATGTCGCTCATATGCA 636
|||||
Db 33286 AACATGCTAAGAGATTAACCTTAAGCTTAATTTTAAAGGAAATTAAGCTTCTGTC 33207

QY 637 CAGTTTCGGAGCGCTTAACCTTTATTCAGAGGACTATCAATTCATCAGATATTTGTC 696
|||||
Db 33206 CAGTTTCCTCATCAACCAAAAGAGTATATAAAAAAATCAATTAATCGCAGATATTTGTG 33147

QY 697 AAAAAAAAAAAGACTAA 714
|||||
Db 33146 ATGAAAAAAAAAAAAA 33129
```

RESULT 13

AC021710 195932 bp DNA linear HTG 04-APR-2000
LOCUS AC021710
DEFINITION Homo sapiens chromosome X clone RP11-12D5 map X, WORKING DRAFT
SEQUENCE, 9 unordered pieces.
AC021710
AC021710
AC021710.4 GI:7408916
VERSION AC021710.4 GI:7408916
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 195932)
AUTHORS Bliren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome X, clone RP11-12D5

JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 195932)
Birten, B., Linton, L., Nusbbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barra, N., Beckert, R., Beda, F.,
Boguslavsky, L., Bouknight, B., Brown, A., Burkett, G., Castle, A.,
Choepe, Y., Collangelo, M., Collins, S., Collamore, A., Cooke, P.,
DeArillano, K., Dewar, K., Domino, M., Doyle, M., Fennestor, J.,
Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,
Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGuck, A., McKernan, K.,
McPheeters, R., Meldrum, J., Menus, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Strange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A., and Zody, M.

Direct Submission
Submitted (19-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 4, 2000 this sequence version replaced gi:6939568.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE
JOURNAL

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L5924
Center clone name: 12_D_5

Summary Statistics

Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 189137 bases at least Q40
Consensus quality: 192288 bases at least Q30
Consensus quality: 193686 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 195132; sum-of-coverage
Quality coverage: 5.0 in Q20 bases; sum-of-coverage
Quality coverage: 4.9 in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1279: contig of 1279 bp in length
* 1280 1379: gap of 100 bp
* 1380 10400: contig of 9021 bp in length
* 10401 10500: gap of 100 bp
* 10501 19494: contig of 8994 bp in length
* 19495 19594: gap of 100 bp
* 19595 32680: contig of 13086 bp in length
* 32681 32780: gap of 100 bp
* 32781 49150: contig of 16370 bp in length
* 49151 49250: gap of 100 bp
* 49251 73382: contig of 24132 bp in length
* 73383 73482: gap of 100 bp
* 73483 104624: contig of 31142 bp in length
* 104625 104724: gap of 100 bp
* 104725 142393: contig of 37669 bp in length
* 142394 142494: gap of 100 bp
* 142494 195932: contig of 53439 bp in length.
Location/Qualifiers
1. 195932

FEATURES
source

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="X"
/clone="RP11-1205"
/clone_1b="RP11-11 Human Male BAC"
1. 1279
/note="assembly_fragment"
1380. 10400
/note="assembly_fragment"
10501. 19494
/note="assembly_fragment"
19595. 32680
/note="assembly_fragment"
32781. 49150
/note="assembly_fragment"
clone_end: 7
vector_side: right"
49251. 73382
/note="assembly_fragment"
73483. 104624
/note="assembly_fragment"
104725. 142393
/note="assembly_fragment"
142494. 195932
/note="assembly_fragment"
clone_end: 8P6
vector_side: left"

ORIGIN

Query Match 5.9%; Score 42.8; DB 2; Length 195932;
Best Local Similarity 51.0%; Pred. No. 0.61; Mismatches 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 517 GAACATAGTGAATTAAGGATGTAACCTTCGATGAGAAATTGACGAGAAAACTAT 576
DB 188972 GAACATAGTGAATTAAGGATGTAACCTTCGATGAGAAATTGACGATGTTTGGAAATTA 189031
QY 577 GGCTAGCTGGAGGTTGTTTCAATCATATATTAAGGAGAAATTGCTCACTATGTA 636
DB 189032 AACATGCTTAAGGATTAACCTTAAGTGTATTTTAAGGAAATTAAGGAGCTTCGTGTC 189091
QY 637 CAGTTCTGCGAGCGCTTAACCTTTATTCGACGAGCTCAATCATGATATTTGTC 696
DB 189092 CAGTTCTCTCATCAACCAAGAGATATTAAGGAGATTAATTAATGCGCATCATGTCG 189151
QY 697 AAAAAAAAAAAGACTTA 714
DB 189152 ATGAAAAAAAAAAAAA 189169

RESULT 14
AL359542/c 28858 bp DNA linear PRI 09-MAR-2001

LOCUS Human DNA sequence from clone RP6-190D15 on chromosome Xq25-26.1,
DEFINITION complete sequence.
ACCESSION AL359542
VERSION AL359542.13 GI:13274757

KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 28858)

AUTHORS Heath, P.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-2001) Sanger Centre, Hinxton, Cambridgehire,
CB10 1SA, UK. E-mail enquiries: hunquerry@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk

COMMENT On Mar 12, 2001 this sequence version replaced gi:13092292.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/chx> RP6-190015 is from the library RPCT-6 constructed by the group of Pieker de Jong. For further details see <http://www.choiri.org/bacpac/home.htm>

IMPORTANT: This sequence is not the entire insert of clone RP6-190D15. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP4-537K3 is at 28759 in this sequence. The true right end of clone RP4-5339 is at 100 in this sequence.

FEATURES

```

source
1. .28858
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="q25-26.1"
/clone="RP6-190D15"
/clone_lib="RPCT-6"
39. .338
repeat_region
/note="AluY repeat: matches 1. .300 of consensus"
347. .646
repeat_region
/note="AluSX repeat: matches 1. .299 of consensus"
3076. .3129
/note="27 copies 2 mer tc 77% conserved"
5914. .6224
repeat_region
/note="AluYb8 repeat: matches 1. .311 of consensus"
6287. .6596
/note="AluSq repeat: matches 1. .312 of consensus"
6597. .8469
repeat_region
/note="L1MB8 repeat: matches 4267. .6133 of consensus"
8470. .8757
repeat_region
/note="AluSx repeat: matches 5. .302 of consensus"
8758. .8796
repeat_region
/note="L1MB8 repeat: matches 6133. .6171 of consensus"
8818. .9104
repeat_region
/note="AluSg1 repeat: matches 1. .289 of consensus"
9345. .9465
repeat_region
/note="L2 repeat: matches 2602. .2728 of consensus"
10287. .10362
repeat_region
/note="38 copies 2 mer aa 69% conserved"
10380. .10462
repeat_region
/note="L1MB3 repeat: matches 6035. .6127 of consensus"
10464. .10518
repeat_region
/note="78K repeat: matches 1. .53 of consensus"
10835. .11133
repeat_region
/note="AluSX repeat: matches 1. .298 of consensus"
11889. .12080
repeat_region
/note="AluSG repeat: matches 24. .296 of consensus"
12081. .12122
repeat_region
/note="AluSg repeat: matches 253. .294 of consensus"
12927. .12974
repeat_region
/note="24 copies 2 mer gt 72% conserved"
13691. .13842

```


This Page Blank (uspto)